(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/03793 A2

(51) International Patent Classification[7]: A01K 67/027

(21) International Application Number: PCT/US01/21923

(22) International Filing Date: 10 July 2001 (10.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/217,058	10 July 2000 (10.07.2000)	US
60/217,179	10 July 2000 (10.07.2000)	US
60/217,223	10 July 2000 (10.07.2000)	US
60/217,253	10 July 2000 (10.07.2000)	US
60/217,255	10 July 2000 (10.07.2000)	US
60/217,256	10 July 2000 (10.07.2000)	US
60/217,257	10 July 2000 (10.07.2000)	US
60/217,347	11 July 2000 (11.07.2000)	US
60/217,629	11 July 2000 (11.07.2000)	US
60/217,537	12 July 2000 (12.07.2000)	US
60/218,069	12 July 2000 (12.07.2000)	US
60/218,074	12 July 2000 (12.07.2000)	US
60/218,358	12 July 2000 (12.07.2000)	US
60/221,483	27 July 2000 (27.07.2000)	US
60/223,120	7 August 2000 (07.08.2000)	US
60/223,122	7 August 2000 (07.08.2000)	US
60/243,958	26 October 2000 (26.10.2000)	US
60/249,408	15 November 2000 (15.11.2000)	US
60/252,299	20 November 2000 (20.11.2000)	US
60/262,113	16 January 2001 (16.01.2001)	US
60/262,205	16 January 2001 (16.01.2001)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/217,179 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,058 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,223 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,253 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,255 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,256 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,257 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,347 (CON)
Filed on	11 July 2000 (11.07.2000)
US	60/217,629 (CON)
Filed on	11 July 2000 (11.07.2000)
US	60/217,537 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,069 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,074 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,358 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/221,483 (CON)
Filed on	27 July 2000 (27.07.2000)
US	60/223,120 (CON)
Filed on	7 August 2000 (07.08.2000)
US	60/223,122 (CON)
Filed on	7 August 2000 (07.08.2000)
US	60/243,958 (CON)
Filed on	26 October 2000 (26.10.2000)
US	60/249,408 (CON)
Filed on	15 November 2000 (15.11.2000)
US	60/252,299 (CON)
Filed on	20 November 2000 (20.11.2000)
US	60/262,113 (CON)
Filed on	16 January 2001 (16.01.2001)
US	60/262,205 (CON)
Filed on	16 January 2001 (16.01.2001)

(71) Applicant *(for all designated States except US)*: **DELTAGEN, INC.** [US/US]; 1003 Hamilton Avenue, Menlo Park, CA 94025 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ALLEN, Keith, D.** [US/US]; 224 Custer Trail, Cary, NC 27513 (US). **BRENNAN, Thomas, J.** [US/US]; 325 Rockwood Drive, South San Francisco, CA 94080 (US).

(74) Agents: **BURKE, John, E.** et al.; Deltagen, Inc., 1003 Hamilton Avenue, Menlo Park, CA 94025 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,

[Continued on next page]

(54) Title: TRANSGENIC MICE CONTAINING TARGETED GENE DISRUPTIONS

(57) Abstract: The present invention relates to transgenic animals, as well as compositions and methods relating to the characterization of gene function. Specifically, the present invention provides transgenic mice comprising mutations in a GPCR gene. Such transgenic mice are useful as models for disease and for identifying agents that modulate gene expression and gene function, and as potential treatments for various disease states and disease conditions.

WO 02/03793 A2

LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TRANSGENIC MICE CONTAINING TARGETED GENE DISRUPTIONS

Field of the Invention

The present invention relates to transgenic animals, compositions and methods relating to the characterization of gene function.

Background of the Invention

Many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers such as cAMP. The membrane protein gene superfamily of G-protein coupled receptors (GPCRs) include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors. GPCRs have been characterized as having seven putative transmembrane domains (designated TM1, TM2, TM3, TM4, TM5, TM6, and TM7), which are believed to represent transmembrane α-helices connected by extracellular or cytoplasmic loops. Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters. Different G-protein α-subunits preferentially stimulate particular effectors to modulate various biological functions in a cell.

Over the past 15 years, nearly 350 therapeutic agents targeting 7 transmembrane receptors have been successfully introduced onto the market. As these receptors have an established, proven history as therapeutic targets, a clear need exists for identification and characterization of GPCRs which can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

Given the importance of GPCRs, a clear need exists for identification and characterization of GPCRs which can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

Summary of the Invention

The present invention generally relates to transgenic animals, as well as to compositions and methods relating to the characterization of gene function.

The present invention provides transgenic cells comprising a disruption in a targeted gene. The transgenic cells of the present invention are comprised of any cells capable of undergoing homologous recombination. Preferably, the cells of the present invention are stem cells and more preferably, embryonic stem (ES) cells, and most preferably, murine ES cells. According to one embodiment, the transgenic cells are produced by introducing a targeting construct into a stem cell to produce a homologous recombinant, resulting in a mutation of the targeted gene. In another embodiment, the transgenic cells are derived from the transgenic animals described below. The cells derived from the transgenic animals includes cells that are isolated or present in a tissue or organ, and any cell lines or any progeny thereof.

The present invention also provides a targeting construct and methods of producing the targeting construct that when introduced into stem cells produces a homologous recombinant. In one embodiment, the targeting construct of the present invention comprises first and second polynucleotide sequences that are homologous to the targeted gene. The targeting construct also comprises a polynucleotide sequence that encodes a selectable marker that is preferably positioned between the two different homologous polynucleotide sequences in the construct. The targeting construct may also comprise other regulatory elements that may enhance homologous recombination.

The present invention further provides non-human transgenic animals and methods of producing such non-human transgenic animals comprising a disruption in a GPCR gene. The transgenic animals of the present invention include transgenic animals that are heterozygous and homozygous for a mutation in the GPCR gene. In one aspect, the transgenic animals of the present invention are defective in the function of the GPCR gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a GPCR gene.

The transgenic animals of the present invention include transgenic animals that are heterozygous and homozygous for a mutation in the melanocortin-3 receptor gene. In one aspect, the transgenic animals of the present invention are defective in the function of the melanocortin-3 receptor gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a melanocortin-3 receptor gene. In a preferred embodiment, the non-human transgenic animals with a mutation in the melanocortin-3 receptor gene comprise abnormalities in the kidney. In another preferred embodiment, the melanocortin-3 receptor mutants demonstrate unilateral renal agenesis. In yet another preferred embodiment, the melanocortin-3 receptor mutants demonstrate passive, hypoactive behavior.

The transgenic animals of the present invention also include transgenic animals that are heterozygous and homozygous for a mutation in the 5-HT-2B gene. In one aspect, the transgenic animals of the present invention are defective in the function of the 5-HT-2B gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a 5-HT-2B gene. In a preferred embodiment, the non-human transgenic animals of the present invention comprise abnormalities in embryonic development. In another preferred embodiment, the non-human transgenic animals of the present invention die between embryonic day 8.5 and 9.5.

The present invention further includes transgenic animals that are heterozygous and homozygous for a mutation in the chemokine receptor 9A gene. In one aspect, the transgenic animals of the present invention are defective in the function of the chemokine receptor 9A gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a chemokine receptor 9A gene. In a preferred embodiment, the non-human transgenic

animals of the present invention demonstrate a decrease in agility, coordination or balance as compared to normal animals.

The present invention also includes transgenic animals that are heterozygous and homozygous for a mutation in the glucocorticoid-induced receptor gene. In one aspect, the transgenic animals of the present invention are defective in the function of the glucocorticoid-induced receptor gene. In another aspect, the transgenic animals of the present invention comprise a phenotype associated with having a mutation in a glucocorticoid-induced receptor gene. In another preferred embodiment, the non-human transgenic animals of the present invention demonstrate hyperactivity and decreased anxiety. In yet another preferred embodiment, the non-human transgenic animals of the present invention demonstrate decreased propensity towards behavioral despair or depression.

The present invention also provides methods of identifying agents capable of affecting a phenotype of a transgenic animal. For example, a putative agent is administered to the transgenic animal and a response of the transgenic animal to the putative agent is measured and compared to the response of a "normal" or wild type mouse, or alternatively compared to a transgenic animal control (without agent administration). The invention further provides agents identified according to such methods. The present invention also provides methods of identifying agents useful as therapeutic agents for treating conditions associated with a disruption of the GPCR gene.

The present invention further provides a method of identifying agents having an effect on GPCR expression or function. The method includes administering an effective amount of the agent to a transgenic animal, preferably a mouse. The method includes measuring a response of the transgenic animal, for example, to the agent, and comparing the response of the transgenic animal to a control animal, which may be, for example, a wild-type animal or alternatively, a transgenic animal control. Compounds that may have an effect on GPCR expression or function may also be screened against cells in cell-based assays, for example, to identify such compounds.

The invention also provides cell lines comprising nucleic acid sequences of a GPCR gene. Such cell lines may be capable of expressing such sequences by virtue of operable linkage to a promoter functional in the cell line. Preferably, expression of the GPCR gene sequence is under the control of an inducible promoter. Also provided are methods of identifying agents that interact with the GPCR gene, comprising the steps of contacting the GPCR gene with an agent and detecting an agent/GPCR gene complex. Such complexes can be detected by, for example, measuring expression of an operably linked detectable marker.

The invention further provides methods of treating diseases or conditions associated with a disruption in a GPCR gene, and more particularly, to a disruption in the expression or function of the GPCR gene. In a preferred embodiment, methods of the present invention involve treating diseases or conditions associated with a disruption in the GPCR gene's expression or function, including

administering to a subject in need, a therapeutic agent that effects GPCR expression or function. In accordance with this embodiment, the method comprises administration of a therapeutically effective amount of a natural, synthetic, semi-synthetic, or recombinant GPCR gene, GPCR gene products or fragments thereof as well as natural, synthetic, semi-synthetic or recombinant analogs.

The present invention further provides methods of treating diseases or conditions associated with disrupted targeted gene expression or function, wherein the methods comprise detecting and replacing through gene therapy mutated GPCR genes.

Definitions

The term "gene" refers to (a) a gene containing at least one of the DNA sequences disclosed herein; (b) any DNA sequence that encodes the amino acid sequence encoded by the DNA sequences disclosed herein and/or; (c) any DNA sequence that hybridizes to the complement of the coding sequences disclosed herein. Preferably, the term includes coding as well as noncoding regions, and preferably includes all sequences necessary for normal gene expression including promoters, enhancers and other regulatory sequences.

The terms "polynucleotide" and "nucleic acid molecule" are used interchangeably to refer to polymeric forms of nucleotides of any length. The polynucleotides may contain deoxyribonucleotides, ribonucleotides and/or their analogs. Nucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The term "polynucleotide" includes single-, double-stranded and triple helical molecules.

"Oligonucleotide" refers to polynucleotides of between 5 and about 100 nucleotides of single- or double-stranded DNA. Oligonucleotides are also known as oligomers or oligos and may be isolated from genes, or chemically synthesized by methods known in the art. A "primer" refers to an oligonucleotide, usually single-stranded, that provides a 3'-hydroxyl end for the initiation of enzyme-mediated nucleic acid synthesis. The following are non-limiting embodiments of polynucleotides: a gene or gene fragment, exons, introns, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes and primers. A nucleic acid molecule may also comprise modified nucleic acid molecules, such as methylated nucleic acid molecules and nucleic acid molecule analogs. Analogs of purines and pyrimidines are known in the art, and include, but are not limited to, aziridinycytosine, 4-acetylcytosine, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethyl-aminomethyluracil, inosine, N6-isopentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, pseudouracil, 5-pentylnyluracil and 2,6-diaminopurine. The use of uracil as a substitute for thymine in a deoxyribonucleic acid is also considered an analogous form of pyrimidine.

A "fragment" of a polynucleotide is a polynucleotide comprised of at least 9 contiguous nucleotides, preferably at least 15 contiguous nucleotides and more preferably at least 45 nucleotides, of coding or non-coding sequences.

The term "gene targeting" refers to a type of homologous recombination that occurs when a fragment of genomic DNA is introduced into a mammalian cell and that fragment locates and recombines with endogenous homologous sequences.

The term "homologous recombination" refers to the exchange of DNA fragments between two DNA molecules or chromatids at the site of homologous nucleotide sequences.

The term "homologous" as used herein denotes a characteristic of a DNA sequence having at least about 70 percent sequence identity as compared to a reference sequence, typically at least about 85 percent sequence identity, preferably at least about 95 percent sequence identity, and more preferably about 98 percent sequence identity, and most preferably about 100 percent sequence identity as compared to a reference sequence. Homology can be determined using a "BLASTN" algorithm. It is understood that homologous sequences can accommodate insertions, deletions and substitutions in the nucleotide sequence. Thus, linear sequences of nucleotides can be essentially identical even if some of the nucleotide residues do not precisely correspond or align. The reference sequence may be a subset of a larger sequence, such as a portion of a gene or flanking sequence, or a repetitive portion of a chromosome.

The term "target gene" (alternatively referred to as "target gene sequence" or "target DNA sequence" or "target sequence") refers to any nucleic acid molecule or polynucleotide of any gene to be modified by homologous recombination. The target sequence includes an intact gene, an exon or intron, a regulatory sequence or any region between genes. The target gene comprises a portion of a particular gene or genetic locus in the individual's genomic DNA. As provided herein, the target gene of the present invention is a GPCR gene. A "GPCR gene" refers to: a melanocortin-3 receptor gene, 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, or an adenosine 3 receptor gene.

A "melanocortin-3 receptor gene" refers to a sequence comprising SEQ ID NO:1 or comprising the sequence encoding the melanocortin-3 receptor [identified in Genbank as Accession No.: X74983; GI NO: 400473]. In one aspect, the coding sequence of the melanocortin-3 receptor gene comprises SEQ ID NO:1 or comprises the melanocortin-3 receptor gene identified in Genbank as Accession No.: X74983; GI NO: 400473.

A "5-HT-2B gene" refers to a sequence comprising SEQ ID NO:5 or comprising the sequence encoding the 5-HT-2B [identified in Genebank as Accession No.: Z15119; GI NO: 54089]. In one

aspect, the coding sequence of the 5-HT-2B gene comprises SEQ ID NO:5 or comprises the 5-HT-2B gene identified in Genebank as Accession No.: Z15119; GI NO: 54089.

A "chemokine receptor 9A gene" refers to a sequence comprising SEQ ID NO:9 or comprising the sequence encoding the chemokine receptor 9A [identified in Genbank as Accession No.: U45982; GI NO: 1245054]. In one aspect, the coding sequence of the chemokine receptor 9A gene comprises SEQ ID NO:9 or comprises the chemokine receptor 9A gene identified in Genbank as Accession No.: U45982; GI NO: 1245054.

A "glucocorticoid-induced receptor gene" refers to a sequence comprising SEQ ID NO:1 or comprising the sequence encoding the glucocorticoid-induced receptor [identified in Genbank as Accession No.: M80481; GI NO: 193516]. In one aspect, the coding sequence of the glucocorticoid-induced receptor gene comprises SEQ ID NO:1 or comprises the glucocorticoid-induced receptor gene identified in Genbank as Accession No.: M80481; GI NO: 193516.

An "orphan GPR10 (UHR-1) gene" refers to a sequence comprising SEQ ID NO:17 or comprising the sequence encoding the orphan GPR10 (UHR-1) [identified in Genbank as Accession No.: S77867; GI NO: 998527]. In one aspect, the coding sequence of the orphan GPR10 (UHR-1) gene comprises SEQ ID NO:17 or comprises the orphan GPR10 (UHR-1) gene identified in Genbank as Accession No.: S77867; GI NO: 998527.

An "orphan GPR14 gene" refers to a sequence comprising SEQ ID NO:21 or comprising the sequence encoding the orphan GPR14 [identified in Genbank as Accession No.: U32673; GI NO: 1002742]. In one aspect, the coding sequence of the orphan GPR14 gene comprises SEQ ID NO:21 or comprises the orphan GPR14 gene identified in Genbank as Accession No.: U32673; GI NO: 1002742.

An "orphan GPR15 gene" refers to a sequence comprising SEQ ID NO:25 or comprising the sequence encoding the orphan GPR15 receptor [identified in Genbank as Accession No.: U34806; GI NO: 1171145]. In one aspect, the coding sequence of the orphan GPR15 gene comprises SEQ ID NO:25 or comprises the orphan GPR15 gene identified in Genbank as Accession No.: U34806; GI NO: 1171145.

A "beta chemokine receptor (E01) gene" refers to a sequence comprising SEQ ID NO:29 or comprising the sequence encoding the beta chemokine receptor (E01) receptor [identified in Genbank as Accession No.: AF030185; GI NO: 2623640]. In one aspect, the coding sequence of the beta chemokine receptor (E01) gene comprises SEQ ID NO:29 or comprises the beta chemokine receptor (E01) gene identified in Genbank as Accession No.: AF030185; GI NO: 2623640.

A "endothelial differentiation GPCR 3 (EDG3) gene" refers to a sequence comprising SEQ ID NO:33 or comprising the sequence encoding the endothelial differentiation GPCR 3 (EDG3) [identified in Genbank as Accession No.: NM010101; GI NO: 6753715]. In one aspect, the coding

sequence of the endothelial differentiation GPCR 3 (EDG3) gene comprises SEQ ID NO:33 or comprises the endothelial differentiation GPCR 3 (EDG3) gene identified in Genbank as Accession No.: NM010101; GI NO: 6753715.

An "ATP receptor P2U1gene" refers to a sequence comprising SEQ ID NO:37 or comprising the sequence encoding the ATP receptor P2U1 [identified in Genbank as Accession No.: L14751; GI NO: 309457]. In one aspect, the coding sequence of the ATP receptor P2U1 gene comprises SEQ ID NO:37 or comprises the ATP receptor P2U1 identified in Genbank as Accession No.: L14751; GI NO: 309457.

An "adenosine 3 receptor gene" refers to a sequence comprising SEQ ID NO:41 or comprising the sequence encoding the adenosine 3 receptor [identified in Genbank as Accession No.: L20331; GI NO: 438796]. In one aspect, the coding sequence of the adenosine 3 receptor gene comprises SEQ ID NO:41 or comprises the adenosine 3 receptor identified in Genbank as Accession No.: L20331; GI NO: 438796.

"Disruption" of a GPCR gene occurs when a fragment of genomic DNA locates and recombines with an endogenous homologous sequence. These sequence disruptions or modifications may include insertions, missense, frameshift, deletion, or substitutions, or replacements of DNA sequence, or any combination thereof. Insertions include the insertion of entire genes, which may be of animal, plant, fungal, insect, prokaryotic, or viral origin. Disruption, for example, can alter or replace a promoter, enhancer, or splice site of a GPCR gene, and can alter the normal gene product by inhibiting its production partially or completely or by enhancing the normal gene product's activity.

The term, "transgenic cell", refers to a cell containing within its genome a GPCR gene that has been disrupted, modified, altered, or replaced completely or partially by the method of gene targeting.

The term "transgenic animal" refers to an animal that contains within its genome a specific gene that has been disrupted by the method of gene targeting. The transgenic animal includes both the heterozygote animal (*i.e.*, one defective allele and one wild-type allele) and the homozygous animal (*i.e.*, two defective alleles).

As used herein, the terms "selectable marker" or "positive selection marker" refers to a gene encoding a product that enables only the cells that carry the gene to survive and/or grow under certain conditions. For example, plant and animal cells that express the introduced neomycin resistance (*Neo*^{r}) gene are resistant to the compound G418. Cells that do not carry the *Neo*^{r} gene marker are killed by G418. Other positive selection markers will be known to those of skill in the art.

A "host cell" includes an individual cell or cell culture that can be or has been a recipient for vector(s) or for incorporation of nucleic acid molecules and/or proteins. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in

total DNA complement) to the original parent due to natural, accidental, or deliberate mutation. A host cell includes cells transfected with the constructs of the present invention.

The term "modulates" as used herein refers to the inhibition, reduction, increase or enhancement of a GPCR function, expression, activity, or alternatively a phenotype associated with a disruption in a GPCR gene.

The term "ameliorates" refers to a decreasing, reducing or eliminating of a condition, disease, disorder, or phenotype, including an abnormality or symptom associated with a disruption in a GPCR gene.

The term "abnormality" refers to any disease, disorder, condition, or phenotype in which a disruption of a GPCR gene is implicated, including pathological conditions.

Brief Description of the Drawings

Figure 1 shows the polynucleotide sequence for a GPCR (melanocortin-3 receptor gene; SEQ ID NO:1). Figure 1 also shows the amino acid sequence for the melanocortin-3 receptor gene (SEQ ID NO:2).

Figure 2 (Panels A and B) shows design of the targeting construct used to disrupt the melanocortin-3 receptor gene. Figure 2 (Panel B) shows the sequences identified as SEQ ID NO:3 and SEQ ID NO:4, which were used as the targeting arms (homologous sequences) in the melanocortin-3 receptor gene targeting construct.

Figure 3 shows the polynucleotide sequence for a 5-HT-2B (SEQ ID NO:5). Figure 3 also shows the amino acid sequence for the 5-HT-2B (SEQ ID NO:6).

Figure 4 (Panels A and B) shows design of the targeting construct used to disrupt 5-HT-2B genes. Figure 4 (Panel B) shows the sequences identified as SEQ ID NO:7 and SEQ ID NO:8, which were used as the targeting arms (homologous sequences) in the 5-HT-2B targeting construct.

Figure 5 shows the polynucleotide sequence for an chemokine receptor 9A (SEQ ID NO:9). Figure 5 also shows the amino acid sequence for the chemokine receptor 9A (SEQ ID NO:10).

Figure 6 (Panels A and B) shows design of the targeting construct used to disrupt chemokine receptor 9A genes. Figure 6 (Panel B) shows the sequences identified as SEQ ID NO:11 and SEQ ID NO:12, which were used as the targeting arms (homologous sequences) in the chemokine receptor 9A targeting construct.

Figure 7 shows the polynucleotide sequence for an glucocorticoid-induced receptor (SEQ ID NO:13). Figure 7 also shows the amino acid sequence for the glucocorticoid-induced receptor (SEQ ID NO:14).

Figure 8 (Panels A and B) shows design of the targeting construct used to disrupt glucocorticoid-induced receptor genes. Figure 8 (Panel B) shows the sequences identified as SEQ ID

NO:15 and SEQ ID NO:16, which were used as the targeting arms (homologous sequences) in the glucocorticoid-induced receptor targeting construct.

Figure 9 shows the polynucleotide sequence for a orphan GPR10 (UHR-1); SEQ ID NO:17). Figure 9 also shows the amino acid sequence for the orphan GPR10 (UHR-1) gene (SEQ ID NO:18).

Figure 10 (Panels A and B) shows design of the targeting construct used to disrupt the orphan GPR10 (UHR-1) gene. Figure 10 (Panel B) shows the sequences identified as SEQ ID NO:19 and SEQ ID NO:20, which were used as the targeting arms (homologous sequences) in the orphan GPR10 (UHR-1) gene targeting construct.

Figure 11 shows the polynucleotide sequence for a GPCR (orphan GPR14 gene; SEQ ID NO:21). Figure 11 also shows the amino acid sequence for the orphan GPR14 gene (SEQ ID NO:22).

Figure 12 (Panels A and B) shows design of the targeting construct used to disrupt the orphan GPR14 gene. Figure 12 (Panel B) shows the sequences identified as SEQ ID NO:23 and SEQ ID NO:24, which were used as the targeting arms (homologous sequences) in the orphan GPR14 gene targeting construct.

Figure 13 shows the polynucleotide sequence for a GPCR (orphan GPR15 gene; SEQ ID NO:25). Figure 13 also shows the amino acid sequence for the orphan GPR15 gene (SEQ ID NO:26).

Figure 14 (Panels A and B) shows design of the targeting construct used to disrupt the orphan GPR15 gene. Figure 14 (Panel B) shows the sequences identified as SEQ ID NO:27 and SEQ ID NO:28, which were used as the targeting arms (homologous sequences) in the orphan GPR15 gene targeting construct.

Figure 15 shows the polynucleotide sequence for a GPCR (beta chemokine receptor (E01) gene; SEQ ID NO:29). Figure 15 also shows the amino acid sequence for the beta chemokine receptor (E01) gene (SEQ ID NO:30).

Figure 16 (Panels A and B) shows design of the targeting construct used to disrupt the beta chemokine receptor (E01) gene. Figure 16 (Panel B) shows the sequences identified as SEQ ID NO:31 and SEQ ID NO:32, which were used as the targeting arms (homologous sequences) in the beta chemokine receptor (E01) gene targeting construct.

Figure 17 shows the polynucleotide sequence for a GPCR (endothelial differentiation GPCR 3 (EDG3) gene; SEQ ID NO:33). Figure 17 also shows the amino acid sequence for the endothelial differentiation GPCR 3 (EDG3) gene (SEQ ID NO:34).

Figure 18 (Panels A and B) shows design of the targeting construct used to disrupt the endothelial differentiation GPCR 3 (EDG3) gene. Figure 18 (Panel B) shows the sequences identified as SEQ ID NO:35 and SEQ ID NO:36, which were used as the targeting arms (homologous sequences) in the endothelial differentiation GPCR 3 (EDG3) gene targeting construct.

Figure 19 shows the polynucleotide sequence for a GPCR (ATP receptor P2U1 gene; SEQ ID NO:1). Figure 19 also shows the amino acid sequence for the ATP receptor P2U1 gene (SEQ ID NO:38).

Figure 20 (Panels A and B) shows design of the targeting construct used to disrupt the ATP receptor P2U1 gene. Figure 20 (Panel B) shows the sequences identified as SEQ ID NO:39 and SEQ ID NO:40, which were used as the targeting arms (homologous sequences) in the ATP receptor P2U1 gene targeting construct.

Figure 21 shows the polynucleotide sequence for a GPCR (adenosine 3 receptor gene; SEQ ID NO:41). Figure 21 also shows the amino acid sequence for the adenosine 3 receptor gene (SEQ ID NO:42).

Figure 22 (Panels A and B) shows design of the targeting construct used to disrupt the adenosine 3 receptor gene. Figure 22 (Panel B) shows the sequences identified as SEQ ID NO:43 and SEQ ID NO:44, which were used as the targeting arms (homologous sequences) in the adenosine 3 receptor gene targeting construct.

Detailed Description of the Invention

The invention is based, in part, on the evaluation of the expression and role of genes and gene expression products, primarily those associated with a GPCR. Among others, the invention permits the definition of disease pathways and the identification of diagnostically and therapeutically useful targets. For example, genes that are mutated or down-regulated under disease conditions may be involved in causing or exacerbating the disease condition. Treatments directed at up-regulating the activity of such genes or treatments that involve alternate pathways, may ameliorate the disease condition.

Generation of Targeting Construct

The targeting construct of the present invention may be produced using standard methods known in the art. (*see, e.g.*, Sambrook, *et al.*, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; E.N. Glover (eds.), 1985, *DNA Cloning: A Practical Approach*, Volumes I and II; M.J. Gait (ed.), 1984, *Oligonucleotide Synthesis*; B.D. Hames & S.J. Higgins (eds.), 1985, Nucleic Acid Hybridization; B.D. Hames & S.J. Higgins (eds.), 1984, *Transcription and Translation*; R.I. Freshney (ed.), 1986, Animal Cell Culture; Immobilized Cells and Enzymes, IRL Press, 1986; B. Perbal, 1984, A Practical Guide To Molecular Cloning; F.M. Ausubel *et al.*, 1994, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc.). For example, the targeting construct may be prepared in accordance with conventional ways, where sequences may be synthesized, isolated from natural sources, manipulated, cloned, ligated, subjected to *in vitro* mutagenesis, primer repair, or the like. At various stages, the joined sequences may be cloned, and analyzed by restriction analysis, sequencing, or the like.

The targeting DNA can be constructed using techniques well known in the art. For example, the targeting DNA may be produced by chemical synthesis of oligonucleotides, nick-translation of a double-stranded DNA template, polymerase chain-reaction amplification of a sequence (or ligase chain reaction amplification), purification of prokaryotic or target cloning vectors harboring a sequence of interest (*e.g.*, a cloned cDNA or genomic DNA, synthetic DNA or from any of the afore-mentioned combination) such as plasmids, phagemids, YACs, cosmids, bacteriophage DNA, other viral DNA or replication intermediates, or purified restriction fragments thereof, as well as other sources of single and double-stranded polynucleotides having a desired nucleotide sequence. More-over, the length of homology may be selected using known methods in the art. For example, selection may be based on the sequence composition and complexity of the predetermined endogenous target DNA sequence(s).

The targeting construct of the present invention typically comprises a first sequence homologous to a portion or region of the GPCR gene and a second sequence homologous to a second portion or region of the GPCR gene. The targeting construct further comprises a positive selection marker, which is preferably positioned in between the first and the second DNA sequence that are homologous to a portion or region of the target DNA sequence. The positive selection marker may be operatively linked to a promoter and a polyadenylation signal.

Other regulatory sequences known in the art may be incorporated into the targeting construct to disrupt or control expression of a particular gene in a specific cell type. In addition, the targeting construct may also include a sequence coding for a screening marker, for example, green fluorescent protein (GFP), or another modified fluorescent protein.

Although the size of the homologous sequence is not critical and can range from as few as 50 base pairs to as many as 100 kb, preferably each fragment is greater than about 1 kb in length, more preferably between about 1 and about 10 kb, and even more preferably between about 1 and about 5 kb. One of skill in the art will recognize that although larger fragments may increase the number of homologous recombination events in ES cells, larger fragments will also be more difficult to clone.

In a preferred embodiment of the present invention, the targeting construct is prepared directly from a plasmid genomic library using the methods described in pending U.S. Patent Application Serial No. 08/971,310, filed November 17, 1997, the disclosure of which is incorporated herein in its entirety. Generally, a sequence of interest is identified and isolated from a plasmid library in a single step using, for example, long-range PCR. Following isolation of this sequence, a second polynucleotide that will disrupt the target sequence can be readily inserted between two regions encoding the sequence of interest. In accordance with this aspect, the construct is generated in two steps by (1) amplifying (for example, using long-range PCR) sequences homologous to the target sequence, and (2) inserting another polynucleotide (for example a selectable marker) into the PCR

product so that it is flanked by the homologous sequences. Typically, the vector is a plasmid from a plasmid genomic library. The completed construct is also typically a circular plasmid.

In another embodiment, the targeting construct is designed in accordance with the regulated positive selection method described in U.S. Patent Application Ser. No. 60/232,957, filed September 15, 2000, the disclosure of which is incorporated herein in its entirety. The targeting construct is designed to include a PGK-*neo* fusion gene having two *lacO* sites, positioned in the PGK promoter and an NLS-*lacI* gene comprising a lac repressor fused to sequences encoding the NLS from the SV40 T antigen.

In another embodiment, the targeting construct may contain more than one selectable maker gene, including a negative selectable marker, such as the herpes simplex virus tk (HSV-tk) gene. The negative selectable marker may be operatively linked to a promoter and a polyadenylation signal. (*see, e.g.*, U.S. Patent No. 5,464,764; U.S. Patent No. 5,487,992; U.S. Patent No. 5,627,059; and U.S. Patent No. 5,631,153).

Generation of Cells and Confirmation of Homologous Recombination Events

Once an appropriate targeting construct has been prepared, the targeting construct may be introduced into an appropriate host cell using any method known in the art. Various techniques may be employed in the present invention, including, for example, pronuclear microinjection; retrovirus mediated gene transfer into germ lines; gene targeting in embryonic stem cells; electroporation of embryos; sperm-mediated gene transfer; and calcium phosphate/DNA co-precipitates, microinjection of DNA into the nucleus, bacterial protoplast fusion with intact cells, transfection, polycations, *e.g.*, polybrene, polyornithine, *etc.*, or the like (*see, e.g.*, U.S. Patent No. 4,873,191; Van der Putten, *et al.*, 1985, *Proc. Natl. Acad. Sci., USA* 82:6148-6152; Thompson, *et al.*, 1989, *Cell* 56:313-321; Lo, 1983, *Mol Cell. Biol.* 3:1803-1814; Lavitrano, *et al.*,1989, *Cell,* 57:717-723). Various techniques for transforming mammalian cells are known in the art (*see, e.g.*, Gordon, 1989, *Intl. Rev. Cytol.*, 115:171-229; Keown *et al.*, 1989, *Methods in Enzymology*; Keown *et al.*, 1990, *Methods and Enzymology*, Vol. 185, pp. 527-537; Mansour *et al.*, 1988, *Nature*, 336:348-352).

In a preferred aspect of the present invention, the targeting construct is introduced into host cells by electroporation. In this process, electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the construct. The pores created during electroporation permit the uptake of macromolecules such as DNA (*see, e.g.*, Potter, H., *et al.*, 1984, *Proc. Nat'l. Acad. Sci. U.S.A.* 81:7161-7165).

Any cell type capable of homologous recombination may be used in the practice of the present invention. Examples of such target cells include cells derived from vertebrates including mammals such as humans, bovine species, ovine species, murine species, simian species, and ether eucaryotic organisms such as filamentous fungi, and higher multicellular organisms such as plants.

Preferred cell types include embryonic stem (ES) cells, which are typically obtained from pre-implantation embryos cultured *in vitro* (*see, e.g.*, Evans, M. J., *et al.*, 1981, *Nature* 292:154-156; Bradley, M. O., *et al.*, 1984, *Nature* 309:255-258; Gossler *et al.*, 1986, *Proc. Natl. Acad. Sci. USA* 83:9065-9069; and Robertson, *et al.*, 1986, *Nature* 322:445-448). The ES cells are cultured and prepared for introduction of the targeting construct using methods well known to the skilled artisan (*see, e.g.*, Robertson, E. J. ed. "Teratocarcinomas and Embryonic Stem Cells, a Practical Approach", IRL Press, Washington D.C., 1987; Bradley *et al.*, 1986, *Current Topics in Devel. Biol.* 20:357-371; by Hogan *et al., in* "Manipulating the Mouse Embryo": A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor N.Y., 1986; Thomas *et al.*, 1987, *Cell* 51:503; Koller *et al.*, 1991, *Proc. Natl. Acad. Sci. USA*, 88:10730; Dorin *et al.*, 1992, *Transgenic Res.* 1:101; and Veis *et al.*, 1993, *Cell* 75:229). The ES cells that will be inserted with the targeting construct are derived from an embryo or blastocyst of the same species as the developing embryo into which they are to be introduced. ES cells are typically selected for their ability to integrate into the inner cell mass and contribute to the germ line of an individual when introduced into the mammal in an embryo at the blastocyst stage of development. Thus, any ES cell line having this capability is suitable for use in the practice of the present invention.

The present invention may also be used to knockout genes in other cell types, such as stem cells. By way of example, stem cells may be myeloid, lymphoid, or neural progenitor and precursor cells. These cells comprising a disruption or knockout of a gene may be particularly useful in the study of GPCR gene function in individual developmental pathways. Stem cells may be derived from any vertebrate species, such as mouse, rat, dog, cat, pig, rabbit, human, non-human primates and the like.

After the targeting construct has been introduced into cells, the cells where successful gene targeting has occurred are identified. Insertion of the targeting construct into the targeted gene is typically detected by identifying cells for expression of the marker gene. In a preferred embodiment, the cells transformed with the targeting construct of the present invention are subjected to treatment with an appropriate agent that selects against cells not expressing the selectable marker. Only those cells expressing the selectable marker gene survive and/or grow under certain conditions. For example, cells that express the introduced neomycin resistance gene are resistant to the compound G418, while cells that do not express the neo gene marker are killed by G418. If the targeting construct also comprises a screening marker such as GFP, homologous recombination can be identified through screening cell colonies under a fluorescent light. Cells that have undergone homologous recombination will have deleted the GFP gene and will not fluoresce.

If a regulated positive selection method is used in identifying homologous recombination events, the targeting construct is designed so that the expression of the selectable marker gene is

regulated in a manner such that expression is inhibited following random integration but is permitted (derepressed) following homologous recombination. More particularly, the transfected cells are screened for expression of the *neo* gene, which requires that (1) the cell was successfully electroporated, and (2) *lac* repressor inhibition of *neo* transcription was relieved by homologous recombination. This method allows for the identification of transfected cells and homologous recombinants to occur in one step with the addition of a single drug.

Alternatively, a positive-negative selection technique may be used to select homologous recombinants. This technique involves a process in which a first drug is added to the cell population, for example, a neomycin-like drug to select for growth of transfected cells, *i.e.* positive selection. A second drug, such as FIAU is subsequently added to kill cells that express the negative selection marker, *i.e.* negative selection. Cells that contain and express the negative selection marker are killed by a selecting agent, whereas cells that do not contain and express the negative selection marker survive. For example, cells with non-homologous insertion of the construct express HSV thymidine kinase and therefore are sensitive to the herpes drugs such as gancyclovir (GANC) or FIAU (1-(2-deoxy 2-fluoro-B-D-arabinofluranosyl)-5-iodouracil) (*see, e.g.*, Mansour *et al.*, *Nature* 336:348-352: (1988); Capecchi, *Science* 244:1288-1292, (1989); Capecchi, *Trends in Genet.* 5:70-76 (1989)).

Successful recombination may be identified by analyzing the DNA of the selected cells to confirm homologous recombination. Various techniques known in the art, such as PCR and/or Southern analysis may be used to confirm homologous recombination events.

Homologous recombination may also be used to disrupt genes in stem cells, and other cell types, which are not totipotent embryonic stem cells. By way of example, stem cells may be myeloid, lymphoid, or neural progenitor and precursor cells. Such transgenic cells may be particularly useful in the study of GPCR gene function in individual developmental pathways. Stem cells may be derived from any vertebrate species, such as mouse, rat, dog, cat, pig, rabbit, human, non-human primates and the like.

In cells that are not totipotent it may be desirable to knock out both copies of the target using methods that are known in the art. For example, cells comprising homologous recombination at a target locus that have been selected for expression of a positive selection marker (*e.g.*, Neo^r) and screened for non-random integration, can be further selected for multiple copies of the selectable marker gene by exposure to elevated levels of the selective agent (*e.g.*, G418). The cells are then analyzed for homozygosity at the target locus. Alternatively, a second construct can be generated with a different positive selection marker inserted between the two homologous sequences. The two constructs can be introduced into the cell either sequentially or simultaneously, followed by appropriate selection for each of the positive marker genes. The final cell is screened for homologous recombination of both alleles of the target.

Production of Transgenic Animals

Selected cells are then injected into a blastocyst (or other stage of development suitable for the purposes of creating a viable animal, such as, for example, a morula) of an animal (*e.g.*, a mouse) to form chimeras (*see e.g.*, Bradley, A. *in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed., IRL, Oxford, pp. 113-152 (1987)). Alternatively, selected ES cells can be allowed to aggregate with dissociated mouse embryo cells to form the aggregation chimera. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Chimeric progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA. In one embodiment, chimeric progeny mice are used to generate a mouse with a heterozygous disruption in the GPCR gene. Heterozygous transgenic mice can then be mated. It is well know in the art that typically ¼ of the offspring of such matings will have a homozygous disruption in the GPCR gene.

The heterozygous and homozygous transgenic mice can then be compared to normal, wild type mice to determine whether disruption of the GPCR gene causes phenotypic changes, especially pathological changes. For example, heterozygous and homozygous mice may be evaluated for phenotypic changes by physical examination, necropsy, histology, clinical chemistry, complete blood count, body weight, organ weights, and cytological evaluation of bone marrow.

In one embodiment, the phenotype (or phenotypic change) associated with a disruption in the GPCR gene is placed into or stored in a database. Preferably, the database includes: (i) genotypic data (*e.g.*, identification of the disrupted gene) and (ii) phenotypic data (*e.g.*, phenotype(s) resulting from the gene disruption) associated with the genotypic data. The database is preferably electronic. In addition, the database is preferably combined with a search tool so that the database is searchable.

Conditional Transgenic Animals

The present invention further contemplates conditional transgenic or knockout animals, such as those produced using recombination methods. Bacteriophage P1 Cre recombinase and flp recombinase from yeast plasmids are two non-limiting examples of site-specific DNA recombinase enzymes that cleave DNA at specific target sites (lox P sites for cre recombinase and frt sites for flp recombinase) and catalyze a ligation of this DNA to a second cleaved site. A large number of suitable alternative site-specific recombinases have been described, and their genes can be used in accordance with the method of the present invention. Such recombinases include the Int recombinase of bacteriophage λ (with or without Xis) (Weisberg, R. *et al.*, in *Lambda II*, (Hendrix, R., *et al.*, Eds.), Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 211-50 (1983), herein incorporated by reference); TpnI and the β-lactamase transposons (Mercier, *et al.*, *J. Bacteriol.*, 172:3745-57 (1990)); the Tn3 resolvase (Flanagan & Fennewald *J. Molec. Biol.*, 206:295-304 (1989); Stark, *et al.*, *Cell*,

58:779-90 (1989)); the yeast recombinases (Matsuzaki, *et al., J. Bacteriol.*, 172:610-18 (1990)); the *B. subtilis* SpoIVC recombinase (Sato, *et al., J. Bacteriol.* 172:1092-98 (1990)); the Flp recombinase (Schwartz & Sadowski, *J. Molec.Biol.*, 205:647-658 (1989); Parsons, *et al., J. Biol. Chem.*, 265:4527-33 (1990); Golic & Lindquist, *Cell*, 59:499-509 (1989); Amin, *et al., J. Molec. Biol.*, 214:55-72 (1990)); the Hin recombinase (Glasgow, *et al., J. Biol. Chem.*, 264:10072-82 (1989)); immunoglobulin recombinases (Malynn, *et al., Cell*, 54:453-460 (1988)); and the Cin recombinase (Haffter & Bickle, *EMBO J.*, 7:3991-3996 (1988); Hubner, *et al., J. Molec. Biol.*, 205:493-500 (1989)), all herein incorporated by reference. Such systems are discussed by Echols (*J. Biol. Chem.* 265:14697-14700 (1990)); de Villartay (*Nature*, 335:170-74 (1988)); Craig, (*Ann. Rev. Genet.*, 22:77-105 (1988)); Poyart-Salmeron, *et al.*, (*EMBO J.* 8:2425-33 (1989)); Hunger-Bertling, *et al.*,(*Mol Cell. Biochem.*, 92:107-16 (1990)); and Cregg & Madden (*Mol. Gen. Genet.*, 219:320-23 (1989)), all herein incorporated by reference.

Cre has been purified to homogeneity, and its reaction with the loxP site has been extensively characterized (Abremski & Hess *J. Mol. Biol.* 259:1509-14 (1984), herein incorporated by reference). Cre protein has a molecular weight of 35,000 and can be obtained commercially from New England Nuclear/DuPont. The cre gene (which encodes the Cre protein) has been cloned and expressed (Abremski, *et al., Cell* 32:1301-11 (1983), herein incorporated by reference). The Cre protein mediates recombination between two loxP sequences (Sternberg, *et al., Cold Spring Harbor Symp. Quant. Biol.* 45:297-309 (1981)), which may be present on the same or different DNA molecule. Because the internal spacer sequence of the loxP site is asymmetrical, two loxP sites can exhibit directionality relative to one another (Hoess & Abremski *Proc. Natl. Acad. Sci. U.S.A.* 81:1026-29 (1984)). Thus, when two sites on the same DNA molecule are in a directly repeated orientation, Cre will excise the DNA between the sites (Abremski, *et al., Cell* 32:1301-11 (1983)). However, if the sites are inverted with respect to each other, the DNA between them is not excised after recombination but is simply inverted. Thus, a circular DNA molecule having two loxP sites in direct orientation will recombine to produce two smaller circles, whereas circular molecules having two loxP sites in an inverted orientation simply invert the DNA sequences flanked by the loxP sites. In addition, recombinase action can result in reciprocal exchange of regions distal to the target site when targets are present on separate DNA molecules.

Recombinases have important application for characterizing gene function in knockout models. When the constructs described herein are used to disrupt GPCR genes, a fusion transcript can be produced when insertion of the positive selection marker occurs downstream (3') of the translation initiation site of the GPCR gene. The fusion transcript could result in some level of protein expression with unknown consequence. It has been suggested that insertion of a positive selection marker gene can affect the expression of nearby genes. These effects may make it difficult to determine gene

function after a knockout event since one could not discern whether a given phenotype is associated with the inactivation of a gene, or the transcription of nearby genes. Both potential problems are solved by exploiting recombinase activity. When the positive selection marker is flanked by recombinase sites in the same orientation, the addition of the corresponding recombinase will result in the removal of the positive selection marker. In this way, effects caused by the positive selection marker or expression of fusion transcripts are avoided.

In one embodiment, purified recombinase enzyme is provided to the cell by direct microinjection. In another embodiment, recombinase is expressed from a co-transfected construct or vector in which the recombinase gene is operably linked to a functional promoter. An additional aspect of this embodiment is the use of tissue-specific or inducible recombinase constructs that allow the choice of when and where recombination occurs. One method for practicing the inducible forms of recombinase-mediated recombination involves the use of vectors that use inducible or tissue-specific promoters or other gene regulatory elements to express the desired recombinase activity. The inducible expression elements are preferably operatively positioned to allow the inducible control or activation of expression of the desired recombinase activity. Examples of such inducible promoters or other gene regulatory elements include, but are not limited to, tetracycline, metallothionine, ecdysone, and other steroid-responsive promoters, rapamycin responsive promoters, and the like (No, *et al.*, *Proc. Natl. Acad. Sci. USA*, 93:3346-51 (1996); Furth, *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:9302-6 (1994)). Additional control elements that can be used include promoters requiring specific transcription factors such as viral, promoters. Vectors incorporating such promoters would only express recombinase activity in cells that express the necessary transcription factors.

Models for Disease

The cell- and animal-based systems described herein can be utilized as models for diseases. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate disease animal models. In addition, cells from humans may be used. These systems may be used in a variety of applications. Such assays may be utilized as part of screening strategies designed to identify agents, such as compounds that are capable of ameliorating disease symptoms. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions that may be effective in treating disease.

Cell-based systems may be used to identify compounds that may act to ameliorate disease symptoms. For example, such cell systems may be exposed to a compound suspected of exhibiting an ability to ameliorate disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of disease symptoms in the exposed cells. After exposure, the cells are

examined to determine whether one or more of the disease cellular phenotypes has been altered to resemble a more normal or more wild type, non-disease phenotype.

In addition, animal-based disease systems, such as those described herein, may be used to identify compounds capable of ameliorating disease symptoms. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions that may be effective in treating a disease or other phenotypic characteristic of the animal. For example, animal models may be exposed to a compound or agent suspected of exhibiting an ability to ameliorate disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of disease symptoms in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of disorders associated with the disease. Exposure may involve treating mother animals during gestation of the model animals described herein, thereby exposing embryos or fetuses to the compound or agent that may prevent or ameliorate the disease or phenotype. Neonatal, juvenile, and adult animals can also be exposed.

More particularly, using the animal models of the invention, specifically, transgenic mice, methods of identifying agents, including compounds are provided, preferably, on the basis of the ability to affect at least one phenotype associated with a disruption in a GPCR gene. In one embodiment, the present invention provides a method of identifying agents having an effect on GPCR expression or function. The method includes measuring a physiological response of the animal, for example, to the agent, and comparing the physiological response of such animal to a control animal, wherein the physiological response of the animal comprising a disruption in a GPCR as compared to the control animal indicates the specificity of the agent. A "physiological response" is any biological or physical parameter of an animal that can be measured. Molecular assays (*e.g.*, gene transcription, protein production and degradation rates), physical parameters (*e.g.*, exercise physiology tests, measurement of various parameters of respiration, measurement of heart rate or blood pressure, measurement of bleeding time, aPTT.T, or TT), and cellular assays (*e.g.*,. immunohistochemical assays of cell surface markers, or the ability of cells to aggregate or proliferate) can be used to assess a physiological response.

The transgenic animals and cells of the present invention may be utilized as models for diseases, disorders, or conditions associated with phenotypes relating to a disruption in a GPCR. The present invention also provides a unique animal model for testing and developing new treatments relating to the behavioral phenotypes. Analysis of the behavioral phenotype allows for the development of an animal model useful for testing, for instance, the efficacy of proposed genetic and pharmacological therapies for human genetic diseases, such as neurological, neuropsychological, or psychotic illnesses.

A statistical analysis of the various behaviors measured can be carried out using any conventional statistical program routinely used by those skilled in the art (such as, for example, "Analysis of Variance" or ANOVA). A "p" value of about 0.05 or less is generally considered to be statistically significant, although slightly higher p values may still be indicative of statistically significant differences. To statistically analyze abnormal behavior, a comparison is made between the behavior of a transgenic animal (or a group thereof) to the behavior of a wild-type mouse (or a group thereof), typically under certain prescribed conditions. "Abnormal behavior" as used herein refers to behavior exhibited by an animal having a disruption in the GPCR gene, *e.g.* transgenic animal, which differs from an animal without a disruption in the GPCR gene, *e.g.* wild-type mouse. Abnormal behavior consists of any number of standard behaviors that can be objectively measured (or observed) and compared. In the case of comparison, it is preferred that the change be statistically significant to confirm that there is indeed a meaningful behavioral difference between the knockout animal and the wild-type control animal. Examples of behaviors that may be measured or observed include, but are not limited to, ataxia, rapid limb movement, eye movement, breathing, motor activity, cognition, emotional behaviors, social behaviors, hyperactivity, hypersensitivity, anxiety, impaired learning, abnormal reward behavior, and abnormal social interaction, such as aggression.

A series of tests may be used to measure the behavioral phenotype of the animal models of the present invention, including neurological and neuropsychological tests to identify abnormal behavior. These tests may be used to measure abnormal behavior relating to, for example, learning and memory, eating, pain, aggression, sexual reproduction, anxiety, depression, schizophrenia, and drug abuse. (*see, e.g.*, Crawley & Paylor, *Hormones and Behavior* 31:197-211 (1997)).

The social interaction test involves exposing a mouse to other animals in a variety of settings. The social behaviors of the animals (*e.g.*, touching, climbing, sniffing, and mating) are subsequently evaluated. Differences in behaviors can then be statistically analyzed and compared (*see, e.g.*, S. E. File, *et al.*, *Pharmacol. Bioch. Behav.* 22:941-944 (1985); R. R. Holson, *Phys. Behav.* 37:239-247 (1986)). Examplary behavioral tests include the following.

The mouse startle response test typically involves exposing the animal to a sensory (typically auditory) stimulus and measuring the startle response of the animal (*see, e.g.*, M. A. Geyer, *et al.*, *Brain Res. Bull.* 25:485-498 (1990); Paylor and Crawley, *Psychopharmacology* 132:169-180 (1997)). A pre-pulse inhibition test can also be used, in which the percent inhibition (from a normal startle response) is measured by "cueing" the animal first with a brief low-intensity pre-pulse prior to the startle pulse.

The electric shock test generally involves exposure to an electrified surface and measurement of subsequent behaviors such as, for example, motor activity, learning, social behaviors. The behaviors are measured and statistically analyzed using standard statistical tests. (*see, e.g.*, G. J. Kant, *et al.*,

Pharm. Bioch. Behav. 20:793-797 (1984); N. J. Leidenheimer, *et al.*, *Pharmacol. Bioch. Behav.* 30:351-355 (1988)).

The tail-pinch or immobilization test involves applying pressure to the tail of the animal and/or restraining the animal's movements. Motor activity, social behavior, and cognitive behavior are examples of the areas that are measured. (*see, e.g.*, M. Bertolucci D'Angic, *et al.*, *Neurochem.* 55:1208-1214 (1990)).

The novelty test generally comprises exposure to a novel environment and/or novel objects. The animal's motor behavior in the novel environment and/or around the novel object are measured and statistically analyzed. (*see, e.g.*, D. K. Reinstein, *et al.*, *Pharm. Bioch. Behav.* 17:193-202 (1982); B. Poucet, *Behav. Neurosci.* 103:1009-10016 (1989); R. R. Holson, *et al.*, *Phys. Behav.* 37:231-238 (1986)). This test may be used to detect visual processing deficiencies or defects.

The learned helplessness test involves exposure to stresses, for example, noxious stimuli, which cannot be affected by the animal's behavior. The animal's behavior can be statistically analyzed using various standard statistical tests. (*see, e.g.*, A. Leshner, *et al.*, *Behav. Neural Biol.* 26:497-501 (1979)).

Alternatively, a tail suspension test may be used, in which the "immobile" time of the mouse is measured when suspended "upside-down" by its tail. This is a measure of whether the animal struggles, an indicator of depression. In humans, depression is believed to result from feelings of a lack of control over one's life or situation. It is believed that a depressive state can be elicited in animals by repeatedly subjecting them to aversive situations over which they have no control. A condition of "learned helplessness" is eventually reached, in which the animal will stop trying to change its circumstances and simply accept its fate. Animals that stop struggling sooner are believed to be more prone to depression. Studies have shown that the administration of certain antidepressant drugs prior to testing increases the amount of time that animals struggle before giving up.

The Morris water-maze test comprises learning spatial orientations in water and subsequently measuring the animal's behaviors, such as, for example, by counting the number of incorrect choices. The behaviors measured are statistically analyzed using standard statistical tests. (*see, e.g.*, E. M. Spruijt, *et al.*, *Brain Res.* 527:192-197 (1990)).

Alternatively, a Y-shaped maze may be used (*see, e.g.*, McFarland, D.J., *Pharmacology, Biochemistry and Behavior* 32:723-726 (1989); Dellu, F., *et al.*, *Neurobiology of Learning and Memory* 73:31-48 (2000)). The Y-maze is generally believed to be a test of cognitive ability. The dimensions of each arm of the Y-maze can be, for example, approximately 40 cm x 8 cm x 20 cm, although other dimensions may be used. Each arm can also have, for example, sixteen equally spaced photobeams to automatically detect movement within the arms. At least two different tests can be performed using such a Y-maze. In a continuous Y-maze paradigm, mice are allowed to explore all

three arms of a Y-maze for, *e.g.*, approximately 10 minutes. The animals are continuously tracked using photobeam detection grids, and the data can be used to measure spontaneous alteration and positive bias behavior. Spontaneous alteration refers to the natural tendency of a "normal" animal to visit the least familiar arm of a maze. An alternation is scored when the animal makes two consecutive turns in the same direction, thus representing a sequence of visits to the least recently entered arm of the maze. Position bias determines egocentrically defined responses by measuring the animal's tendency to favor turning in one direction over another. Therefore, the test can detect differences in an animal's ability to navigate on the basis of allocentric or egocentric mechanisms. The two-trial Y-maze memory test measures response to novelty and spatial memory based on a free-choice exploration paradigm. During the first trial (acquisition), the animals are allowed to freely visit two arms of the Y-maze for, *e.g.*, approximately 15 minutes. The third arm is blocked off during this trial. The second trial (retrieval) is performed after an intertrial interval of, *e.g.*, approximately 2 hours. During the retrieval trial, the blocked arm is opened and the animal is allowed access to all three arms for, *e.g.*, approximately 5 minutes. Data are collected during the retrieval trial and analyzed for the number and duration of visits to each arm. Because the three arms of the maze are virtually identical, discrimination between novelty and familiarity is dependent on "environmental" spatial cues around the room relative to the position of each arm. Changes in arm entry and duration of time spent in the novel arm in a transgenic animal model may be indicative of a role of that gene in mediating novelty and recognition processes.

The passive avoidance or shuttle box test generally involves exposure to two or more environments, one of which is noxious, providing a choice to be learned by the animal. Behavioral measures include, for example, response latency, number of correct responses, and consistency of response. (*see, e.g.*, R. Ader, *et al.*, *Psychon. Sci.* 26:125-128 (1972); R. R. Holson, *Phys. Behav.* 37:221-230 (1986)). Alternatively, a zero-maze can be used. In a zero-maze, the animals can, for example, be placed in a closed quadrant of an elevated annular platform having, *e.g.*, 2 open and 2 closed quadrants, and are allowed to explore for approximately 5 minutes. This paradigm exploits an approach-avoidance conflict between normal exploratory activity and an aversion to open spaces in rodents. This test measures anxiety levels and can be used to evaluate the effectiveness of anti-anxiolytic drugs. The time spent in open quadrants versus closed quadrants may be recorded automatically, with, for example, the placement of photobeams at each transition site.

The food avoidance test involves exposure to novel food and objectively measuring, for example, food intake and intake latency. The behaviors measured are statistically analyzed using standard statistical tests. (*see, e.g.*, B. A. Campbell, *et al.*, *J. Comp. Physiol. Psychol.* 67:15-22 (1969)).

The elevated plus-maze test comprises exposure to a maze, without sides, on a platform, the animal's behavior is objectively measured by counting the number of maze entries and maze learning. The behavior is statistically analyzed using standard statistical tests. (*see, e.g.*, H. A. Baldwin, *et al.*, *Brain Res. Bull*, 20:603-606 (1988)).

The stimulant-induced hyperactivity test involves injection of stimulant drugs (*e.g.*, amphetamines, cocaine, PCP, and the like), and objectively measuring, for example, motor activity, social interactions, cognitive behavior. The animal's behaviors are statistically analyzed using standard statistical tests. (*see, e.g.*, P. B. S. Clarke, *et al.*, *Psychopharmacology* 96:511-520 (1988); P. Kuczenski, *et al.*, *J. Neuroscience* 11:2703-2712 (1991)).

The self-stimulation test generally comprises providing the mouse with the opportunity to regulate electrical and/or chemical stimuli to its own brain. Behavior is measured by frequency and pattern of self-stimulation. Such behaviors are statistically analyzed using standard statistical tests. (*see, e.g.*, S. Nassif, *et al.*, *Brain Res.*, 332:247-257 (1985); W. L. Isaac, *et al.*, *Behav. Neurosci.* 103:345-355 (1989)).

The reward test involves shaping a variety of behaviors, *e.g.*, motor, cognitive, and social, measuring, for example, rapidity and reliability of behavioral change, and statistically analyzing the behaviors measured. (*see, e.g.*, L. E. Jarrard, *et al.*, Exp. Brain Res. 61:519-530 (1986)).

The DRL (differential reinforcement to low rates of responding) performance test involves exposure to intermittent reward paradigms and measuring the number of proper responses, *e.g.*, lever pressing. Such behavior is statistically analyzed using standard statistical tests. (*see, e.g.*, J. D. Sinden, *et al.*, Behav. Neurosci. 100:320-329 (1986); V. Nalwa, *et al.*, Behav Brain Res. 17:73-76 (1985); and A. J. Nonneman, *et al.*, J. Comp. Physiol. Psych. 95:588-602 (1981)).

The spatial learning test involves exposure to a complex novel environment, measuring the rapidity and extent of spatial learning, and statistically analyzing the behaviors measured. (*see, e.g.*, N. Pitsikas, *et al.*, Pharm. Bioch. Behav. 38:931-934 (1991); B. Poucet, *et al.*, Brain Res. 37:269-280 (1990); D. Christie, *et al.*, Brain Res. 37:263-268 (1990); and F. Van Haaren, *et al.*, Behav. Neurosci. 102:481-488 (1988)). Alternatively, an open-field (of) test may be used, in which the greater distance traveled for a given amount of time is a measure of the activity level and anxiety of the animal. When the open field is a novel environment, it is believed that an approach-avoidance situation is created, in which the animal is "torn" between the drive to explore and the drive to protect itself. Because the chamber is lighted and has no places to hide other than the corners, it is expected that a "normal" mouse will spend more time in the corners and around the periphery than it will in the center where there is no place to hide. "Normal" mice will, however, venture into the central regions as they explore more and more of the chamber. It can then be extrapolated that especially anxious mice will spend most of their time in the corners, with relatively little or no exploration of the central region,

whereas bold (*i.e.*, less anxious) mice will travel a greater distance, showing little preference for the periphery versus the central region.

The visual, somatosensory and auditory neglect tests generally comprise exposure to a sensory stimulus, objectively measuring, for example, orientating responses, and statistically analyzing the behaviors measured. (*see, e.g.*, J. M. Vargo, *et al.*, Exp. Neurol. 102:199-209 (1988)).

The consummatory behavior test generally comprises feeding and drinking, and objectively measuring quantity of consumption. The behavior measured is statistically analyzed using standard statistical tests. (*see, e.g.*, P. J. Fletcher, *et al.*, Psychopharmacol. 102:301-308 (1990); M. G. Corda, *et al.*,, Proc. Nat'l Acad. Sci. USA 80:2072-2076 (1983)).

A visual discrimination test can also be used to evaluate the visual processing of an animal. One or two similar objects are placed in an open field and the animal is allowed to explore for about 5-10 minutes. The time spent exploring each object (proximity to, *i.e.*, movement within, *e.g.*, about 3-5 cm of the object is considered exploration of an object) is recorded. The animal is then removed from the open field, and the objects are replaced by a similar object and a novel object. The animal is returned to the open field and the percent time spent exploring the novel object over the old object is measured (again, over about a 5-10 minute span). "Normal" animals will typically spend a higher percentage of time exploring the novel object rather than the old object. If a delay is imposed between sampling and testing, the memory task becomes more hippocampal-dependent. If no delay is imposed, the task is more based on simple visual discrimination. This test can also be used for olfactory discrimination, in which the objects (preferably, simple blocks) can be sprayed or otherwise treated to hold an odor. This test can also be used to determine if the animal can make gustatory discriminations; animals that return to the previously eaten food instead of novel food exhibit gustatory neophobia.

A hot plate analgesia test can be used to evaluate an animal's sensitivity to heat or painful stimuli. For example, a mouse can be placed on an approximately 55°C hot plate and the mouse's response latency (*e.g.*, time to pick up and lick a hind paw) can be recorded. These responses are not reflexes, but rather "higher" responses requiring cortical involvement. This test may be used to evaluate a nociceptive disorder.

An accelerating rotarod test may be used to measure coordination and balance in mice. Animals can be, for example, placed on a rod that acts like a rotating treadmill (or rolling log). The rotarod can be made to rotate slowly at first and then progressively faster until it reaches a speed of, *e.g.*, approximately 60 rpm. The mice must continually reposition themselves in order to avoid falling off. The animals are preferably tested in at least three trials, a minimum of 20 minutes apart. Those mice that are able to stay on the rod the longest are believed to have better coordination and balance.

A metrazol administration test can be used to screen animals for varying susceptibilities to seizures or similar events. For example, a 5mg/ml solution of metrazol can be infused through the tail vein of a mouse at a rate of, *e.g.*, approximately 0.375 ml/min. The infusion will cause all mice to experience seizures, followed by death. Those mice that enter the seizure stage the soonest are believed to be more prone to seizures. Four distinct physiological stages can be recorded: soon after the start of infusion, the mice will exhibit a noticeable "twitch", followed by a series of seizures, ending in a final tensing of the body known as "tonic extension", which is followed by death.

<u>GPCR Gene Products</u>

The present invention further contemplates use of the GPCR gene sequence to produce GPCR gene products. GPCR gene products may include proteins that represent functionally equivalent gene products. Such an equivalent gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the gene sequences described herein, but which result in a silent change, thus producing a functionally equivalent GPCR gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent", as utilized herein, refers to a protein capable of exhibiting a substantially similar *in vivo* activity as the endogenous gene products encoded by the GPCR gene sequences. Alternatively, when utilized as part of an assay, "functionally equivalent" may refer to peptides capable of interacting with other cellular or extracellular molecules in a manner substantially similar to the way in which the corresponding portion of the endogenous gene product would.

Other protein products useful according to the methods of the invention are peptides derived from or based on the GPCR gene produced by recombinant or synthetic means (derived peptides).

GPCR gene products may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the gene polypeptides and peptides of the invention by expressing nucleic acid encoding gene sequences are described herein. Methods that are well known to those skilled in the art can be used to construct expression vectors containing gene protein coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination (*see, e.g.*, Sambrook, *et al.*, 1989, *supra*, and Ausubel, *et al.*, 1989, supra). Alternatively, RNA capable of encoding gene protein sequences may be chemically

synthesized using, for example, automated synthesizers (*see, e.g.* Oligonucleotide Synthesis: A Practical Approach, Gait, M. J. ed., IRL Press, Oxford (1984)).

A variety of host-expression vector systems may be utilized to express the gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells that may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the gene protein of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing gene protein coding sequences; yeast (*e.g. Saccharomyces, Pichia*) transformed with recombinant yeast expression vectors containing the gene protein coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing gene protein coding sequences; or mammalian cell systems (*e.g.* COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionine promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5 K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the gene protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, *EMBO J.*, 2:1791-94 (1983)), in which the gene protein coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, *Nucleic Acids Res.*, 13:3101-09 (1985); Van Heeke *et al.*, *J. Biol. Chem.*, 264:5503-9 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned GPCR gene protein can be released from the GST moiety.

In a preferred embodiment, full length cDNA sequences are appended with in-frame Bam HI sites at the amino terminus and Eco RI sites at the carboxyl terminus using standard PCR methodologies (Innis, *et al.* (eds) PCR Protocols: A Guide to Methods and Applications, Academic Press, San

Diego (1990)) and ligated into the pGEX-2TK vector (Pharmacia, Uppsala, Sweden). The resulting cDNA construct contains a kinase recognition site at the amino terminus for radioactive labeling and glutathione S-transferase sequences at the carboxyl terminus for affinity purification (Nilsson, *et al.*, *EMBO J.*, 4: 1075-80 (1985); Zabeau *et al.*, *EMBO J.*, 1: 1217-24 (1982)).

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The gene coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (*see*, *e.g.*, Smith, *et al.*, *J. Virol.* 46: 584-93 (1983); U.S. Patent No. 4,745,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the gene coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing gene protein in infected hosts. (*e.g.*, *see* Logan *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:3655-59 (1984)). Specific initiation signals may also be required for efficient translation of inserted gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (*see* Bitter, *et al.*, *Methods in Enzymol.*, 153:516-44 (1987)).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific

mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the gene protein may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells that stably integrate the plasmid into their chromosomes and grow, to form foci, which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the gene protein.

In a preferred embodiment, timing and/or quantity of expression of the recombinant protein can be controlled using an inducible expression construct. Inducible constructs and systems for inducible expression of recombinant proteins will be well known to those skilled in the art. Examples of such inducible promoters or other gene regulatory elements include, but are not limited to, tetracycline, metallothionine, ecdysone, and other steroid-responsive promoters, rapamycin responsive promoters, and the like (No, *et al.*, *Proc. Natl. Acad. Sci. USA*, 93:3346-51 (1996); Furth, *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:9302-6 (1994)). Additional control elements that can be used include promoters requiring specific transcription factors such as viral, particularly HIV, promoters. In one in embodiment, a Tet inducible gene expression system is utilized. (Gossen *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:5547-51 (1992); Gossen, *et al.*, *Science*, 268:1766-69 (1995)). Tet Expression Systems are based on two regulatory elements derived from the tetracycline-resistance operon of the *E. coli* Tn10 transposon—the tetracycline repressor protein (TetR) and the tetracycline operator sequence (*tetO*) to which TetR binds. Using such a system, expression of the recombinant protein is placed under the control of the *tetO* operator sequence and transfected or transformed into a host cell. In the presence of TetR, which is co-transfected into the host cell, expression of the recombinant protein is repressed due to binding of the TetR protein to the *tetO* regulatory element. High-level, regulated gene expression can then be induced in response to varying concentrations of tetracycline (Tc) or Tc

derivatives such as doxycycline (Dox), which compete with *tetO* elements for binding to TetR. Constructs and materials for tet inducible gene expression are available commercially from CLONTECH Laboratories, Inc., Palo Alto, CA.

When used as a component in an assay system, the gene protein may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the gene protein and a test substance. Any of a variety of suitable labeling systems may be used including but not limited to radioisotopes such as ^{125}I; enzyme labeling systems that generate a detectable calorimetric signal or light when exposed to substrate; and fluorescent labels. Where recombinant DNA technology is used to produce the gene protein for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization and/or detection.

Indirect labeling involves the use of a protein, such as a labeled antibody, which specifically binds to the gene product. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library.

<u>Production of Antibodies</u>

Described herein are methods for the production of antibodies capable of specifically recognizing one or more epitopes. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a GPCR gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal GPCR gene activity. Thus, such antibodies may be utilized as part of disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of GPCR gene proteins, or for the presence of abnormal forms of such proteins.

For the production of antibodies, various host animals may be immunized by injection with the GPCR gene, its expression product or a portion thereof. Such host animals may include but are not limited to rabbits, mice, rats, goats and chickens, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as GPCR gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described

above, may be immunized by injection with gene product supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique of Köhler and Milstein, *Nature*, 256:495-7 (1975); and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor, *et al.*, *Immunology Today*, 4:72 (1983); Cote, *et al.*, *Proc. Natl. Acad. Sci. USA*, 80:2026-30 (1983)), and the EBV-hybridoma technique (Cole, *et al.*, in Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., New York, pp. 77-96 (1985)). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. Production of high titers of mAbs *in vivo* makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison, *et al.*, *Proc. Natl. Acad. Sci.*, 81:6851-6855 (1984); Takeda, *et al.*, *Nature*, 314:452-54 (1985)) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, *Science* 242:423-26 (1988); Huston, *et al.*, *Proc. Natl. Acad. Sci. USA*, 85:5879-83 (1988); and Ward, *et al.*, *Nature*, 334:544-46 (1989)) can be adapted to produce gene-single chain antibodies. Single chain antibodies are typically formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments that can be produced by pepsin digestion of the antibody molecule and the Fab fragments that can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse, *et al.*, *Science*, 246:1275-81 (1989)) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

<u>Screening Methods</u>

The present invention may be employed in a process for screening for agents such as agonists, *i.e.* agents that bind to and activate GPCR polypeptides, or antagonists, *i.e.* inhibit the activity or interaction of GPCR polypeptides with its ligand. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free

preparations, chemical libraries, and natural product mixtures as known in the art. Any methods routinely used to identify and screen for agents that can modulate receptors may be used in accordance with the present invention.

The present invention provides methods for identifying and screening for agents that modulate GPCR expression or function. More particularly, cells that contain and express GPCR gene sequences may be used to screen for therapeutic agents. Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC# TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as HUVEC's and bovine aortic endothelial cells (BAEC's); as well as generic mammalian cell lines such as HeLa cells and COS cells, *e.g.*, COS-7 (ATCC# CRL-1651). Further, such cells may include recombinant, transgenic cell lines. For example, the transgenic mice of the invention may be used to generate cell lines, containing one or more cell types involved in a disease, that can be used as cell culture models for that disorder. While cells, tissues, and primary cultures derived from the disease transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For examples of techniques that may be used to derive a continuous cell line from the transgenic animals, see Small, *et al.*, *Mol. Cell Biol.*, 5:642-48 (1985).

GPCR gene sequences may be introduced into, and overexpressed in, the genome of the cell of interest. In order to overexpress a GPCR gene sequence, the coding portion of the GPCR gene sequence may be ligated to a regulatory sequence that is capable of driving gene expression in the cell type of interest. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation. GPCR gene sequences may also be disrupted or underexpressed. Cells having GPCR gene disruptions or underexpressed GPCR gene sequences may be used, for example, to screen for agents capable of affecting alternative pathways that compensate for any loss of function attributable to the disruption or underexpression.

In vitro systems may be designed to identify compounds capable of binding the GPCR gene products. Such compounds may include, but are not limited to, peptides made of D-and/or L-configuration amino acids (in, for example, the form of random peptide libraries; (*see e.g.*, Lam, *et al.*, *Nature*, 354:82-4 (1991)), phosphopeptides (in, for example, the form of random or partially degenerate, directed phosphopeptide libraries; *see*, *e.g.*, Songyang, *et al.*, *Cell*, 72:767-78 (1993)), antibodies, and small organic or inorganic molecules. Compounds identified may be useful, for example, in modulating the activity of GPCR gene proteins, preferably mutant GPCR gene proteins; elaborating the biological function of the GPCR gene protein; or screening for compounds that disrupt normal GPCR gene interactions or themselves disrupt such interactions.

The principle of the assays used to identify compounds that bind to the GPCR gene protein involves preparing a reaction mixture of the GPCR gene protein and the test compound under

conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring the GPCR gene protein or the test substance onto a solid phase and detecting target protein/test substance complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the GPCR gene protein may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

In practice, microtitre plates are conveniently utilized. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for GPCR gene product or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Compounds that are shown to bind to a particular GPCR gene product through one of the methods described above can be further tested for their ability to elicit a biochemical response from the GPCR gene protein. Agonists, antagonists and/or inhibitors of the expression product can be identified utilizing assays well known in the art.

Antisense, Ribozymes, and Antibodies

Other agents that may be used as therapeutics include the GPCR gene, its expression product(s) and functional fragments thereof. Additionally, agents that reduce or inhibit mutant GPCR gene activity may be used to ameliorate disease symptoms. Such agents include antisense, ribozyme,

and triple helix molecules. Techniques for the production and use of such molecules are well known to those of skill in the art.

Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, *e.g.*, between the -10 and +10 regions of the GPCR gene nucleotide sequence of interest, are preferred.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage. The composition of ribozyme molecules must include one or more sequences complementary to the GPCR gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see U.S. Patent No. 5,093,246, which is incorporated by reference herein in its entirety. As such within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences encoding GPCR gene proteins.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the molecule of interest for ribozyme cleavage sites that include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the GPCR gene containing the cleavage site may be evaluated for predicted structural features, such as secondary structure, that may render the oligonucleotide sequence unsuitable. The suitability of candidate sequences may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription should be single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

It is possible that the antisense, ribozyme, and/or triple helix molecules described herein may reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by both normal and mutant GPCR gene alleles. In order to ensure that substantially normal levels of GPCR gene activity are maintained, nucleic acid molecules that encode and express GPCR gene polypeptides exhibiting normal activity may be introduced into cells that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, it may be preferable to coadminister normal GPCR gene protein into the cell or tissue in order to maintain the requisite level of cellular or tissue GPCR gene activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Various well-known modifications to the DNA molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

Antibodies that are both specific for GPCR gene protein, and in particular, mutant gene protein, and interfere with its activity may be used to inhibit mutant GPCR gene function. Such antibodies may be generated against the proteins themselves or against peptides corresponding to portions of the proteins using standard techniques known in the art and as also described herein. Such antibodies include but are not limited to polyclonal, monoclonal, Fab fragments, single chain antibodies, chimeric antibodies, etc.

In instances where the GPCR gene protein is intracellular and whole antibodies are used, internalizing antibodies may be preferred. However, lipofectin liposomes may be used to deliver the

antibody or a fragment of the Fab region that binds to the GPCR gene epitope into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target or expanded target protein's binding domain is preferred. For example, peptides having an amino acid sequence corresponding to the domain of the variable region of the antibody that binds to the GPCR gene protein may be used. Such peptides may be synthesized chemically or produced via recombinant DNA technology using methods well known in the art (*see, e.g.*, Creighton, Proteins: Structures and Molecular Principles (1984) W.H. Freeman, New York 1983, *supra*; and Sambrook, *et al.*, 1989, supra). Alternatively, single chain neutralizing antibodies that bind to intracellular GPCR gene epitopes may also be administered. Such single chain antibodies may be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco, *et al., Proc. Natl. Acad. Sci. USA*, 90:7889-93 (1993).

RNA sequences encoding GPCR gene protein may be directly administered to a patient exhibiting disease symptoms, at a concentration sufficient to produce a level of GPCR gene protein such that disease symptoms are ameliorated. Patients may be treated by gene replacement therapy. One or more copies of a normal GPCR gene, or a portion of the gene that directs the production of a normal GPCR gene protein with GPCR gene function, may be inserted into cells using vectors that include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be utilized for the introduction of normal GPCR gene sequences into human cells.

Cells, preferably, autologous cells, containing normal GPCR gene expressing gene sequences may then be introduced or reintroduced into the patient at positions that allow for the amelioration of disease symptoms.

Pharmaceutical Compositions, Effective Dosages, and Routes of Administration

The identified compounds that inhibit target mutant gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to treat or ameliorate the disease. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of the disease.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a

delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral, topical, subcutaneous, intraperitoneal, intravenous, intrapleural, intraoccular, intraarterial, or rectal administration. It is also contemplated that pharmaceutical compositions may be administered with other products that potentiate the activity of the compound and optionally, may include other therapeutic ingredients.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.* gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. Oral ingestion is possibly the easiest method of taking any medication. Such a route of administration, is generally simple and straightforward and is frequently the least inconvenient or unpleasant route of administration from the patient's point of view. However, this involves passing the material through the stomach, which is a hostile environment for many materials, including proteins and other biologically active compositions. As the acidic, hydrolytic and proteolytic environment of the stomach has evolved efficiently to digest proteinaceous materials into amino acids and oligopeptides for subsequent anabolism, it is hardly surprising that very little or any of a wide variety of biologically active proteinaceous material, if simply taken orally, would survive its passage through the stomach to be taken up by the body in the small intestine. The result, is that many proteinaceous medicaments must be taken in through another method, such as parenterally, often by subcutaneous, intramuscular or intravenous injection.

Pharmaceutical compositions may also include various buffers (*e.g.*, Tris, acetate, phosphate), solubilizers (*e.g.*, Tween, Polysorbate), carriers such as human serum albumin, preservatives (thimerosol, benzyl alcohol) and anti-oxidants such as ascorbic acid in order to stabilize pharmaceutical activity. The stabilizing agent may be a detergent, such as tween-20, tween-80, NP-40

or Triton X-100. EBP may also be incorporated into particulate preparations of polymeric compounds for controlled delivery to a patient over an extended period of time. A more extensive survey of components in pharmaceutical compositions is found in Remington's Pharmaceutical Sciences, 18th ed., A. R. Gennaro, ed., Mack Publishing, Easton, Pa. (1990).

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device that may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

Diagnostics

A variety of methods may be employed to diagnose disease conditions associated with the GPCR gene. Specifically, reagents may be used, for example, for the detection of the presence of GPCR gene mutations, or the detection of either over or under expression of GPCR gene mRNA.

According to the diagnostic and prognostic method of the present invention, alteration of the wild-type GPCR gene locus is detected. In addition, the method can be performed by detecting the wild-type GPCR gene locus and confirming the lack of a predisposition or neoplasia. "Alteration of a wild-type gene" encompasses all forms of mutations including deletions, insertions and point mutations in the coding and noncoding regions. Deletions may be of the entire gene or only a portion of the gene. Point mutations may result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those that occur only in certain tissues, *e.g.*, in tumor tissue, and are not inherited in the germline. Germline mutations can be found in any of a body's tissues and are inherited. If only a single allele is somatically mutated, an early neoplastic state may be indicated. However, if both alleles are mutated, then a late neoplastic state may be indicated. The finding of gene mutations thus provides both diagnostic and prognostic information. A GPCR gene allele that is not deleted (*e.g.*, that found on the sister chromosome to a chromosome carrying a GPCR gene deletion) can be screened for other mutations, such as insertions, small deletions, and point mutations. Mutations found in tumor tissues may be linked to decreased expression of the GPCR gene product. However, mutations leading to non-functional gene products may also be linked to a cancerous state. Point mutational events may occur in regulatory regions, such as in the promoter of the gene, leading to loss or diminution of expression of the mRNA. Point mutations may also abolish proper RNA

processing, leading to loss of expression of the GPCR gene product, or a decrease in mRNA stability or translation efficiency.

One test available for detecting mutations in a candidate locus is to directly compare genomic target sequences from cancer patients with those from a control population. Alternatively, one could sequence messenger RNA after amplification, *e.g.*, by PCR, thereby eliminating the necessity of determining the exon structure of the candidate gene. Mutations from cancer patients falling outside the coding region of the GPCR gene can be detected by examining the non-coding regions, such as introns and regulatory sequences near or within the GPCR gene. An early indication that mutations in noncoding regions are important may come from Northern blot experiments that reveal messenger RNA molecules of abnormal size or abundance in cancer patients as compared to control individuals.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific gene nucleic acid or anti-gene antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings, to diagnose patients exhibiting disease symptoms or at risk for developing disease.

Any cell type or tissue in which the gene is expressed may be utilized in the diagnostics described below.

DNA or RNA from the cell type or tissue to be analyzed may easily be isolated using procedures that are well known to those in the art. Diagnostic procedures may also be performed *in situ* directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures (*see*, for example, Nuovo, PCR *In Situ* Hybridization:Protocols and Applications, Raven Press, N.Y. (1992)).

Gene nucleotide sequences, either RNA or DNA, may, for example, be used in hybridization or amplification assays of biological samples to detect disease-related gene structures and expression. Such assays may include, but are not limited to, Southern or Northern analyses, restriction fragment length polymorphism assays, single stranded conformational polymorphism analyses, *in situ* hybridization assays, and polymerase chain reaction analyses. Such analyses may reveal both quantitative aspects of the expression pattern of the gene, and qualitative aspects of the gene expression and/or gene composition. That is, such aspects may include, for example, point mutations, insertions, deletions, chromosomal rearrangements, and/or activation or inactivation of gene expression.

Preferred diagnostic methods for the detection of gene-specific nucleic acid molecules may involve for example, contacting and incubating nucleic acids, derived from the cell type or tissue being analyzed, with one or more labeled nucleic acid reagents under conditions favorable for the specific annealing of these reagents to their complementary sequences within the nucleic acid molecule of interest. Preferably, the lengths of these nucleic acid reagents are at least 9 to 30

nucleotides. After incubation, all non-annealed nucleic acids are removed from the nucleic acid:fingerprint molecule hybrid. The presence of nucleic acids from the fingerprint tissue that have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the tissue or cell type of interest may be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtitre plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic acid reagents are easily removed. Detection of the remaining, annealed, labeled nucleic acid reagents is accomplished using standard techniques well-known to those in the art.

Alternative diagnostic methods for the detection of gene-specific nucleic acid molecules may involve their amplification, *e.g.*, by PCR (the experimental embodiment set forth in Mullis U.S. Patent No. 4,683,202 (1987)), ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. USA*, 88:189-93 (1991)), self sustained sequence replication (Guatelli, *et al.*, *Proc. Natl. Acad. Sci.* USA, 87:1874-78 (1990)), transcriptional amplification system (Kwoh, *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:1173-77 (1989)), Q-Beta Replicase (Lizardi *et al.*, *Bio/Technology*, 6:1197 (1988)), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In one embodiment of such a detection scheme, a cDNA molecule is obtained from an RNA molecule of interest (*e.g.*, by reverse transcription of the RNA molecule into cDNA). Cell types or tissues from which such RNA may be isolated include any tissue in which wild type fingerprint gene is known to be expressed. For example, cell types useful in the detection of the melanocortin-3 receptor gene include, but are not limited to, spermatogenic cells of the seminiferous tubules in the testes; cell types useful in the detection of the5-HT2-C receptor gene include, but are not limited to, testis and eye; cell types useful in the detection of the chemokine receptor 9A gene include, but are not limited to, testis, tongue and esophagus, lung, spleen, thymus, lymph nodes, bone marrow, salivary gland, skeletal muscle and small intestine; cell types useful in the detection of the glucocorticoid-induced receptor gene include, but are not limited to, to brain, pharynx, testis and prostate. A sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (*e.g.*, primers) in the reverse transcription and nucleic acid amplification steps of this method may be chosen from among the gene nucleic acid reagents described herein. The preferred lengths of such nucleic acid reagents are at least 15-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be

visualized by standard ethidium bromide staining or by utilizing any other suitable nucleic acid staining method.

Antibodies directed against wild type or mutant gene peptides may also be used as disease diagnostics and prognostics. Such diagnostic methods, may be used to detect abnormalities in the level of gene protein expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of fingerprint gene protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant fingerprint gene protein relative to the normal fingerprint gene protein.

Protein from the tissue or cell type to be analyzed may easily be detected or isolated using techniques that are well known to those of skill in the art, including but not limited to western blot analysis. For a detailed explanation of methods for carrying out western blot analysis, see Sambrook, *et al.* (1989) *supra*, at Chapter 18. The protein detection and isolation methods employed herein may also be such as those described in Harlow and Lane, for example, (Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988)).

Preferred diagnostic methods for the detection of wild type or mutant gene peptide molecules may involve, for example, immunoassays wherein fingerprint gene peptides are detected by their interaction with an anti-fingerprint gene-specific peptide antibody.

For example, antibodies, or fragments of antibodies useful in the present invention may be used to quantitatively or qualitatively detect the presence of wild type or mutant gene peptides. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred if the fingerprint gene peptides are expressed on the cell surface.

The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of fingerprint gene peptides. *In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the fingerprint gene peptides, but also their distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such *in situ* detection.

Immunoassays for wild type, mutant, or expanded fingerprint gene peptides typically comprise incubating a biological sample, such as a biological fluid, a tissue extract, freshly harvested

cells, or cells that have been incubated in tissue culture, in the presence of a detectably labeled antibody capable of identifying fingerprint gene peptides, and detecting the bound antibody by any of a number of techniques well known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled gene-specific antibody. The solid phase support may then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

The terms "solid phase support or carrier" are intended to encompass any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of a given lot of anti-wild type or -mutant fingerprint gene peptide antibody may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

One of the ways in which the gene peptide-specific antibody can be detectably labeled is by linking the same to an enzyme and using it in an enzyme immunoassay (EIA) (Voller, *Ric Clin Lab*, 8:289-98 (1978) ["The Enzyme Linked Immunosorbent Assay (ELISA)", Diagnostic Horizons 2:1-7, 1978, Microbiological Associates Quarterly Publication, Walkersville, Md.]; Voller, *et al.*, *J. Clin. Pathol.*, 31:507-20 (1978); Butler, *Meth. Enzymol.*, 73:482-523 (1981); Maggio (ed.), Enzyme Immunoassay, CRC Press, Boca Raton, Fla. (1980); Ishikawa, *et al.*, (eds.) Enzyme Immunoassay, Igaku-Shoin, Tokyo (1981)). The enzyme that is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-

glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type, mutant, or expanded peptides through the use of a radioimmunoassay (RIA) (*see, e.g.*, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediamine-tetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Throughout this application, various publications, patents and published patent applications are referred to by an identifying citation. The disclosures of these publications, patents and published

patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

The following examples are intended only to illustrate the present invention and should in no way be construed as limiting the subject invention.

Examples

Example 1: Generation and Analysis of Mice Comprising Melanocortin-3 Receptor Gene Disruptions

To investigate the role of melanocortin-3 receptors, disruptions in melanocortin-3 receptor genes were produced by homologous recombination. Specifically, transgenic mice comprising disruptions in melanocortin-3 receptor genes were created. More particularly, as shown in Figure 2, a melanocortin-3 receptor-specific targeting construct having the ability to disrupt or modify melanocortin-3 receptor genes, specifically comprising SEQ ID NO:1 was created using as the targeting arms (homologous sequences) in the construct, the oligonucleotide sequences identified herein as SEQ ID NO:3 or SEQ ID NO:4.

The targeting construct was introduced into ES cells derived from the 129/Sv-+P+Mgf-SLJ/J mouse substrain to generate chimeric mice. The F1 mice were generated by breeding with C57BL/6 females, and the F2 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females.

The transgenic mice comprising disruptions in melanocortin-3 receptor genes were analyzed for phenotypic changes and expression patterns. The phenotypes associated with a disruption in nuclear receptor genes were determined. The homozygous mice demonstrated at least one of the following phenotypes:

Kidney:

Two homozygous mutant mice (83793, 83779) had unilateral renal agenesis, with only one kidney present at necropsy. One heterozygous mouse (83778) also had unilateral renal agenesis, which is occasionally seen in this strain of mice. Therefore, the presence of only one kidney in homozygous mutants is a possible phenotypic change, although it may be due to spontaneous disease or of nonspecific etiology.

Expression:

Tissues of the transgenic animals were analyzed for expression of the target gene. Organs from one heterozygous male and one heterozygous female were frozen, sectioned (10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase. Nuclear Fast Red was used for counterstaining.

Organs and tissues collected and frozen: brain, sciatic nerve, eye, Harderian glands, thymus, spleen, lymph nodes, bone marrow, aorta, heart, lung, liver, gallbladder, pancreas, kidney, urinary

bladder, trachea, larynx, esophagus, thyroid gland, pituitary gland, adrenal glands, salivary glands, stomach, small and large intestines, tongue, skeletal muscle, skin and reproductive system.

In addition, the brain of the heterozygous female was analyzed for lacZ expression as wholemount. The dissected brain was cut longitudinally, fixed and stained using X-Gal as a substrate for beta-galactosidase. To stop the reaction the brain was washed in PBS and fixed in PBS-buffered formaldehyde.

Wild type control tissues were stained for X-gal to reveal background or signals due to endogenous beta-galactosidase activity. The following tissues show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-galactosidase activity.

Example 2: Generation and Analysis of Mice Comprising 5-HT2-B Gene Disruptions

To investigate the role of 5-HT2-B, disruptions in 5-HT2-B genes were produced by homologous recombination. Specifically, transgenic mice comprising disruptions in 5-HT2-B genes were created. More particularly, as shown in Figure 4, an 5-HT2-B -specific targeting construct having the ability to disrupt or modify 5-HT2-B genes, specifically comprising SEQ ID NO:5 was created using as the targeting arms (homologous sequences) in the construct, the oligonucleotide sequences identified herein as SEQ ID NO:7 or SEQ ID NO:8.

The targeting construct was introduced into ES cells derived from the 129/Sv-+P+Mgf-SLJ/J mouse substrain to generate chimeric mice. The F1 mice were generated by breeding with C57BL/6 females, and the F2 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females.

The transgenic mice comprising disruptions in 5-HT2-B genes were analyzed for phenotypic changes and expression patterns. The phenotypes associated with a disruption in nuclear receptor genes were determined. The homozygous mice demonstrated at least one of the following phenotypes:

Embryonic Lethality. Homozygous mutant embryos die between E8.5 and E9.5

Embryos were isolated at 8.5 to 12.5 days post coitum. Homozygous offspring were detected by PCR at E8.5, but not at later stages. At E8.5 homozygous were abnormal and retarded in development. Further, resorptions were found at E8.5 and all later stages examined. Where genotyping was feasible on partial resorptions, some scored as heterozygotes possibly due to maternal cell infiltration. These data suggest that death is occurring at multiple stages of development.

Embryos were isolated at E8.5 to E12.5. Six litters were examined comprising of 47 embryos, resorptions and partial resorptions, of which 36 were successfully genotyped. These results are summarized below in Table 1.

TABLE 1. 5-HT2-B: Embryonic Development

Litter	Embryonic stage	+/+	+/-	-/-	complete resorption/unknown
1	E8.5 (0-8 somites)	0	3	0	4
2	E8.5 (0-9 somites)	4	2	2	2
3	E8.5 (0-9 somites)	6	2	0	0
4	E9.5	2	6	0	2
5	E12.5	3	1	0	4
6	E12.5	3	2	0	2

Expression:

Tissues of the transgenic animals (heterozygotes) were analyzed for expression of the target gene. Organs from one heterozygous male and one heterozygous female were frozen, sectioned (10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase. Nuclear Fast Red was used for counterstaining.

Organs and tissues collected and frozen: brain, sciatic nerve, eye, Harderian glands, thymus, spleen, lymph nodes, bone marrow, aorta, heart, lung, liver, gallbladder, pancreas, kidney, urinary bladder, trachea, larynx, esophagus, thyroid gland, pituitary gland, adrenal glands, salivary glands, stomach, small and large intestines, tongue, skeletal muscle, skin and reproductive system.

In addition, the brain of the heterozygous female was analyzed for lacZ expression as wholemount. The dissected brain was cut longitudinally, fixed and stained using X-Gal as a substrate for beta-galactosidase. To stop the reaction the brain was washed in PBS and fixed in PBS-buffered formaldehyde.

Wild type control tissues were stained for X-gal to reveal background or signals due to endogenous beta-galactosidase activity. The following tissues show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-galactosidase activity.

The results were as follows:

LacZ (beta-galactosidase) expression was detectable in testis and eye.

Male Reproductive System (Testis): A weak, punctuated staining pattern was observable in the epithelium of seminiferous tubules.

Eye: A medium, punctuated staining pattern was detectable in the retina.

Example 3: Generation and Analysis of Mice Comprising Chemokine Receptor 9A Gene Disruptions

To investigate the role of chemokine receptor 9As, disruptions in chemokine receptor 9A genes were produced by homologous recombination. Specifically, transgenic mice comprising disruptions in chemokine receptor 9A genes were created. More particularly, as shown in Figure 6, a

chemokine receptor 9A-specific targeting construct having the ability to disrupt or modify chemokine receptor 9A genes, specifically comprising SEQ ID NO:9 was created using as the targeting arms (homologous sequences) in the construct, the oligonucleotide sequences identified herein as SEQ ID NO:11 or SEQ ID NO:12.

The targeting construct was introduced into ES cells derived from the 129/Sv-+P+Mgf-SLJ/J mouse substrain to generate chimeric mice. The F1 mice were generated by breeding with C57BL/6 females, and the F2 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females.

The transgenic mice comprising disruptions in chemokine receptor 9A genes were analyzed for phenotypic changes and expression patterns. The phenotypes associated with a disruption in nuclear receptor genes were determined. The homozygous mice demonstrated at least one of the following phenotypes:

Expression.

Total RNA was isolated from the organs or tissues from adult C57Bl/6 wild type mice. RNA was DNaseI treated, and reverse transcribed using random primers. The resulting cDNA was checked for the absence of genomic contamination using primers specific to non-transcribed genomic mouse DNA. cDNAs were balanced for concentration using HPRT primers. RNA transcripts were detectable in lung, spleen, thymus, lymph nodes, bone marrow, salivary gland, skeletal muscle and small intestine.

Tissues of the transgenic animals were analyzed for expression of the target gene. Organs from one heterozygous male and one heterozygous female were frozen, sectioned (10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase. Nuclear Fast Red was used for counterstaining.

Organs and tissues collected and frozen: brain, sciatic nerve, eye, Harderian glands, thymus, spleen, lymph nodes, bone marrow, aorta, heart, lung, liver, gallbladder, pancreas, kidney, urinary bladder, trachea, larynx, esophagus, thyroid gland, pituitary gland, adrenal glands, salivary glands, stomach, small and large intestines, tongue, skeletal muscle, skin and reproductive system.

In addition, the brain of the heterozygous female was analyzed for lacZ expression as wholemount. The dissected brain was cut longitudinally, fixed and stained using X-Gal as a substrate for beta-galactosidase. To stop the reaction the brain was washed in PBS and fixed in PBS-buffered formaldehyde.

Wild type control tissues were stained for X-gal to reveal background or signals due to endogenous beta-galactosidase activity. The following tissues show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas

deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-galactosidase activity.

The results were as follows: LacZ (beta-galactosidase) expression was detectable in testis, tongue and esophagus.

Male reproductive system: Many nuclei in seminiferous tubules of the testis stained weakly.
Esophagus: Strong X-Gal signals were present in several cells in the mucous layer.
Tongue: Epithelial cells of the mucous glands stained moderately.
Behavior:

For behavioral studies, homozygous mice were produced as follows:

The targeting construct described above was introduced into ES cells derived from the 129/SvEv mouse substrain to generate chimeric mice. F1N0 mice were generated by breeding with C57BL/6 females. F2N0 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females. F1N0 heterozygotes were backcrossed to C57BL/6 mice to generate F1N1 heterozygotes. F2N1 homozygous mice were produced by intercrossing F1N1 heterozygous males and females.

The homozygous mice demonstrated the following behavioral phenotypes:

When compared to age- and gender-matched wild-type control mice, homozygous mutant mice exhibited decreased agility, coordination, or balance as characterized by decreased performance on an accelerating rotarod. Mutants fell from the accelerating rotarod at lower speeds across all three trials. By the third trial there was a statistically significant difference between the mutants and wild type mice, as shown below in Table 2.

TABLE 2. Chemokine Receptor 9A - Rotarod Test Fall Speed

Gene	Family	SubFamily	Gene Name	Genotype	F Generation	N Generation	Trial	Value	Std. Err.	Count
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	0	Trial 1	5.96	0.79	12
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	0	Trial 2	7.48	1.07	12
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	0	Trial 3	7.57	0.93	12
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	1	Trial 1	6.62	0.42	13
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	1	Trial 2	8.50	0.97	13
365	GPCR	Chemokine	chemokine receptor 9A	+/+	2	1	Trial 3	11.42	0.66	13
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	0	Trial 1	6.17	0.33	10
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	0	Trial 2	6.38	0.54	10
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	0	Trial 3	6.49	0.70	10

Gene	Family	SubFamily	Gene Name	Genotype	F Generation	N Generation	Trial	Value	Std. Err.	Count
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	1	Trial 1	5.70	0.68	10
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	1	Trial 2	7.13	0.82	10
365	GPCR	Chemokine	chemokine receptor 9A	-/-	2	1	Trial 3	8.31	1.11	10

Example 4: Generation and Analysis of Mice Comprising Glucocorticoid-Induced Receptor Gene Disruptions

To investigate the role of glucocorticoid-induced receptors, disruptions in glucocorticoid-induced receptor genes were produced by homologous recombination. Specifically, transgenic mice comprising disruptions in glucocorticoid-induced receptor genes were created. More particularly, as shown in Figure 8, a glucocorticoid-induced receptor-specific targeting construct having the ability to disrupt or modify glucocorticoid-induced receptor genes, specifically comprising SEQ ID NO:13 was created using as the targeting arms (homologous sequences) in the construct, the oligonucleotide sequences identified herein as SEQ ID NO:15 or SEQ ID NO:16.

The targeting construct was introduced into ES cells derived from the 129/Sv-+P+Mgf-SLJ/J mouse substrain to generate chimeric mice. The F1 mice were generated by breeding with C57BL/6 females, and the F2 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females.

The transgenic mice comprising disruptions in glucocorticoid-induced receptor genes were analyzed for phenotypic changes and expression patterns. The phenotypes associated with a disruption in nuclear receptor genes were determined. The homozygous mice demonstrated at least one of the following phenotypes:

Expression.

Tissues of the transgenic animals were analyzed for expression of the target gene. Organs from one heterozygous male and one heterozygous female were frozen, sectioned (10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase. Nuclear Fast Red was used for counterstaining.

Organs and tissues collected and frozen: brain, sciatic nerve, eye, Harderian glands, thymus, spleen, lymph nodes, bone marrow, aorta, heart, lung, liver, gallbladder, pancreas, kidney, urinary bladder, trachea, larynx, esophagus, thyroid gland, pituitary gland, adrenal glands, salivary glands, stomach, small and large intestines, tongue, skeletal muscle, skin and reproductive system.

In addition, the brain of the heterozygous female was analyzed for lacZ expression as wholemount. The dissected brain was cut longitudinally, fixed and stained using X-Gal as a substrate

for beta-galactosidase. To stop the reaction the brain was washed in PBS and fixed in PBS-buffered formaldehyde.

Wild type control tissues were stained for X-gal to reveal background or signals due to endogenous beta-galactosidase activity. The following tissues show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-galactosidase activity.

The results were as follows:

LacZ expression was detectable in brain, pharynx, testis and prostate.

Brain: In wholemount stained brains strong signals were observed in the cerebrum in thalamus, and surrounding areas including the caudate putamen. Faint lacZ expression was apparent along the cerebellar lobes of the cerebellum. On frozen sections of the heterozygous male few cells stained positive in the caudate putamen, but no lacZ signals were detectable in the cerebellum.

Testis: Weak staining was observed in few cells in the seminiferous tubules.

Prostate: Few epithelial cells stained weakly.

Pharynx: Few cells in the connective tissue showed weak lacZ signals.

Behavior.

For behavioral studies, homozygous mice were produced as follows:

The targeting construct described above was introduced into ES cells derived from the 129/SvEv mouse substrain to generate chimeric mice. F1N0 mice were generated by breeding with C57BL/6 females. F2N0 homozygous mutant mice were produced by intercrossing F1 heterozygous males and females. F1N0 heterozygotes were backcrossed to C57BL/6 mice to generate F1N1 heterozygotes. F2N1 homozygous mice were produced by intercrossing F1N1 heterozygous males and females.

The homozygous mice demonstrated the following behavioral phenotypes:

When compared to age- and gender-matched wild-type control mice, homozygous mutant animals from the N1 generation showed a statistically significant increase in total distance traveled and the percent of time spent in the central region of the test chamber in the open field test, relative to N1 generation wild-type animals (Table 3). This suggests that the mutant mice may be more hyperactive and less anxious than their wild-type littermates. However, N0 generation animals did not display any difference in these measurements.

Table 3. Glucocorticoid-Induced Receptor: Open Field– Session Time in Central Region

Genotype	F Generation	N Generation	Gene	Family	SubFamily	Gene Name	Value	Std. Err.	Count
+/+	2	0	359	GPCR	Orphan GPCR	glucocorticoid induced	38.30	6.56	10

Genotype	F Generation	N Generation	Gene	Family	SubFamily	Gene Name	Value	Std. Err.	Count
						receptor			
+/+	2	1	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	29.79	3.94	11
-/-	2	0	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	39.50	8.47	10
-/-	2	1	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	45.40	4.19	10

Homozygous mutant animals from the N1 generation also spent significantly less time immobile in the tail suspension test relative to wild-types (Table 4). This may indicate that mutant mice have less of a propensity towards behavioral despair or depression than their wild-type littermates (anti-depressive behavior phenotype). However, N0 generation animals did not display any difference in this test.

Table 4. Glucocorticoid-Induced Receptor: Tail Suspension – Total Time Immobile

Genotype	F Generation	N Generation	Gene	Family	SubFamily	Gene Name	Value	Std. Err.	Count
+/+	2	0	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	142.18	23.94	10
+/+	2	1	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	173.05	15.51	11
-/-	2	0	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	114.21	28.99	10
-/-	2	1	359	GPCR	Orphan GPCR	glucocorticoid induced receptor	135.22	8.36	10

The discrepancy in the results observed in the Open Field and Tail Suspension Tests between generations may reflect differences in the background strains used to generate the mice.

As is apparent to one of skill in the art, various modifications of the above embodiments can be made without departing from the spirit and scope of this invention. These modifications and variations are within the scope of this invention.

We claim:

1. A targeting construct comprising:

   (a) a first polynucleotide sequence homologous to a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene;

   (b) a second polynucleotide sequence homologous to the target gene; and

   (c) a selectable marker.

2. The targeting construct of claim 1, wherein the targeting construct further comprises a screening marker.

3. A method of producing a targeting construct for a target gene selected from the group consisting of a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:

   (a) obtaining a first polynucleotide sequence homologous to the target gene;

   (b) obtaining a second polynucleotide sequence homologous to the target gene;

   (c) providing a vector comprising a selectable marker; and

   (d) inserting the first and second sequences into the vector, to produce the targeting construct.

4. A method of producing a targeting construct for a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:

   (a) providing a polynucleotide comprising a first sequence homologous to a first region of the target gene and a second sequence homologous to a second region of the target gene; and

   (b) inserting a positive selection marker between the first and second sequences to form the targeting construct.

5. A cell comprising a disruption in a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene,

a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene.

6. The cell of claim 5, wherein the cell is a murine cell.
7. The cell of claim 6, wherein the murine cell is an embryonic stem cell.
8. A non-human transgenic animal comprising a disruption in a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene.
9. A cell derived from the non-human transgenic animal of claim 8.
10. A method of producing a transgenic mouse, the method comprising:
    (a) introducing the targeting construct of claim 1 into a cell;
    (b) introducing the cell into a blastocyst;
    (c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and
    (d) breeding the chimeric mouse to produce the transgenic mouse.
11. A method of identifying an agent that modulates the expression of a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:
    (a) providing a non-human transgenic animal comprising a disruption in the target gene;
    (b) administering an agent to the non-human transgenic animal; and
    (c) determining whether the expression of the disrupted target gene in the non-human transgenic animal is modulated.
12. A method of identifying an agent that modulates the function of a target gene selected from the group consisting of a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:
    (a) providing a non-human transgenic animal comprising a disruption in the target gene;
    (b) administering an agent to the non-human transgenic animal; and

(c) determining whether the function of the disrupted target gene in the non-human transgenic animal is modulated.

13. A method of identifying an agent that modulates the expression of a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:
    (a) providing a cell comprising a disruption in the target gene;
    (b) contacting the cell with an agent; and
    (c) determining whether expression of the target gene is modulated.
14. A method of identifying an agent that modulates the function of a target gene selected from the group consisting of: a melanocortin-3 gene, a 5-HT-2B gene, a chemokine receptor 9A gene, a glucocorticoid-induced receptor gene, an orphan GPR10 (UHR-1) gene, an orphan GPR14 gene, an orphan GPR15 gene, a beta chemokine receptor (E01) gene, an endothelial differentiation GPCR 3 (EDG3) gene, an ATP receptor P2U1gene, and an adenosine 3 receptor gene, the method comprising:
    (a) providing a cell comprising a disruption in the target gene;
    (b) contacting the cell with an agent; and
    (c) determining whether the function of the target gene is modulated.
15. The method of claim 13 or claim 14, wherein the cell is derived from the non-human transgenic animal of claim 8.
16. An agent identified by the method of claim 11, claim 12, claim 13, or claim 14.
17. A targeting construct comprising:
    (a) a first polynucleotide sequence homologous to a melanocortin-3 receptor gene;
    (b) a second polynucleotide sequence homologous to the melanocortin-3 receptor gene; and
    (c) a selectable marker.
18. The targeting construct of claim 17, wherein the targeting construct further comprises a screening marker.
19. A method of producing a targeting construct, the method comprising:
    (a) providing a first polynucleotide sequence homologous to a melanocortin-3 receptor gene;
    (b) providing a second polynucleotide sequence homologous to the melanocortin-3 receptor;
    (c) providing a selectable marker; and

(d) inserting the first sequence, second sequence, and selectable marker into a vector, to produce the targeting construct.

20. A method of producing a targeting construct, the method comprising:

(a) providing a polynucleotide comprising a first sequence homologous to a first region of a melanocortin-3 receptor gene and a second sequence homologous to a second region of a melanocortin-3 receptor gene;

(b) inserting a positive selection marker in between the first and second sequences to form the targeting construct.

21. A cell comprising a disruption in a melanocortin-3 receptor gene.

22. The cell of claim 21, wherein the cell is a murine cell.

23. The cell of claim 22, wherein the murine cell is an embryonic stem cell.

24. A non-human transgenic animal comprising a disruption in a melanocortin-3 receptor gene.

25. A cell derived from the non-human transgenic animal of claim 24.

26. A method of producing a transgenic mouse comprising a disruption in a melanocortin-3 receptor gene, the method comprising:

(a) introducing the targeting construct of claim 17 into a cell;

(b) introducing the cell into a blastocyst;

(c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

(d) breeding the chimeric mouse to produce the transgenic mouse.

27. A method of identifying an agent that modulates the expression of a melanocortin-3 receptor, the method comprising:

(a) providing a non-human transgenic animal comprising a disruption in a melanocortin-3 receptor gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the expression of melanocortin-3 receptor in the non-human transgenic animal is modulated.

28. A method of identifying an agent that modulates the function of a melanocortin-3 receptor, the method comprising:

(a) providing a non-human transgenic animal comprising a disruption in a melanocortin-3 receptor gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the function of the disrupted melanocortin-3 receptor gene in the non-human transgenic animal is modulated.

29. A method of identifying an agent that modulates the expression of melanocortin-3 receptor, the method comprising:
    (a) providing a cell comprising a disruption in a melanocortin-3 receptor gene;
    (b) contacting the cell with an agent; and
    (c) determining whether expression of the melanocortin-3 receptor is modulated.
30. A method of identifying an agent that modulates the function of a melanocortin-3 receptor gene, the method comprising:
    (a) providing a cell comprising a disruption in a melanocortin-3 receptor gene;
    (b) contacting the cell with an agent; and
    (c) determining whether the function of the melanocortin-3 receptor gene is modulated.
31. The method of claim 29 or claim 30, wherein the cell is derived from the non-human transgenic animal of claim 24.
32. An agent identified by the method of claim 27, claim 28, claim 29, or claim 30.
33. A transgenic mouse comprising a disruption in a melanocortin-3 receptor gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: a kidney abnormality or a behavioral abnormality.
34. The transgenic mouse of claim 33, wherein the kidney abnormality is absence of one kidney.
35. The transgenic mouse of claim 33, wherein the kidney abnormality is reduced size of the kidney relative to a wild-type mouse.
36. The transgenic mouse of claim 33, wherein the kidney comprises unilateral renal agenesis.
37. The transgenic mouse of claim 33, wherein the behavioral abnormality is passivity.
38. The transgenic mouse of claim 33, wherein the behavioral abnormality is hypoactivity.
39. The transgenic mouse of claim 33, wherein the behavioral abnormality is decreased locomotion.
40. The transgenic mouse of claim 33, wherein the behavioral abnormality is a decrease in the attempt to escape while being examined relative to a wild type mouse.
41. The transgenic mouse of claim 33, wherein the behavioral abnormality is absence of any attempt to escape while being examined.
42. The transgenic mouse of claim 33, wherein the behavioral abnormality is observed in males.
43. A method of producing a transgenic mouse comprising a disruption in a melanocortin-3 receptor gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: a kidney abnormality or a behavioral abnormality, the method comprising:
    (a) introducing a melanocortin-3 receptor gene targeting construct into a cell;
    (b) introducing the cell into a blastocyst;

(c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

(d) breeding the chimeric mouse to produce the transgenic mouse comprising a disruption in a melanocortin-3 receptor gene.

44. A transgenic mouse produced by the method of claim 43.

45. A cell derived from the transgenic mouse of claim 33 or claim 44.

46. A method of identifying an agent that ameliorates a phenotype associated with a disruption in a melanocortin-3 receptor gene, the method comprising:

(a) administering an agent to a transgenic mouse comprising a disruption in a melanocortin-3 receptor gene; and

(b) determining whether the agent ameliorates at least one of the following phenotypes: a kidney abnormality or a behavioral abnormality.

47. A method of idêntifying an agent that modulates melanocortin-3 receptor expression, the method comprising:

(a) administering an agent to the transgenic mouse comprising a disruption in a melanocortin-3 receptor gene; and

(b) determining whether the agent modulates melanocortin-3 receptor expression in the transgenic mouse, wherein the agent has an effect on at least one of the following behaviors: passivity, locomotion or attempts to escape while being examined.

48. A method of identifying an agent that modulates a behavior associated with a disruption in a melanocortin-3 receptor gene, the method comprising:

(a) administering an agent to a transgenic mouse comprising a disruption in a melanocortin-3 receptor gene; and

(b) determining whether the agent modulates passivity, locomotion or attempts to escape while being examined.

49. A method of identifying an agent that modulates melanocortin-3 receptor gene function, the method comprising:

(a) providing a cell comprising a disruption in a melanocortin-3 receptor gene;

(b) contacting the cell with an agent; and

(c) determining whether the agent modulates melanocortin-3 receptor gene function, wherein the agent modulates a phenotype associated with a disruption in a melanocortin-3 receptor gene.

50. The method of claim 49, wherein the phenotype comprises at least one of the following: a kidney abnormality or a behavioral abnormality.

51. An agent identified by the method of claim 46, claim 47, claim 48, or claim 49.
52. An agonist or antagonist of a melanocortin-3 receptor.
53. A targeting construct comprising:
    (a) a first polynucleotide sequence homologous to a 5-HT-2B gene;
    (b) a second polynucleotide sequence homologous to the 5-HT-2B gene; and
    (a) a selectable marker.
54. The targeting construct of claim 53, wherein the targeting construct further comprises a screening marker.
55. A method of producing a targeting construct, the method comprising:
    (a) providing a first polynucleotide sequence homologous to a 5-HT-2B gene;
    (b) providing a second polynucleotide sequence homologous to the 5-HT-2B ;
    (c) providing a selectable marker; and
    (d) inserting the first sequence, second sequence, and selectable marker into a vector, to produce the targeting construct.
56. A method of producing a targeting construct, the method comprising:
    (a) providing a polynucleotide comprising a first sequence homologous to a first region of a 5-HT-2B gene and a second sequence homologous to a second region of a 5-HT-2B gene;
    (b) inserting a positive selection marker in between the first and second sequences to form the targeting construct.
57. A cell comprising a disruption in a 5-HT-2B gene.
58. The cell of claim 57, wherein the cell is a murine cell.
59. The cell of claim 58, wherein the murine cell is an embryonic stem cell.
60. A non-human transgenic animal comprising a disruption in a 5-HT-2B gene.
61. A cell derived from the non-human transgenic animal of claim 60.
62. A method of producing a transgenic mouse comprising a disruption in a 5-HT-2B gene, the method comprising:
    (a) introducing the targeting construct of claim 53 into a cell;
    (b) introducing the cell into a blastocyst;
    (c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and
    (d) breeding the chimeric mouse to produce the transgenic mouse.
63. A method of identifying an agent that modulates the expression of a 5-HT-2B, the method comprising:
    (a) providing a non-human transgenic animal comprising a disruption in a 5-HT-2B gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the expression of 5-HT-2B in the non-human transgenic animal is modulated.

64. A method of identifying an agent that modulates the function of a 5-HT-2B, the method comprising:

(a) providing a non-human transgenic animal comprising a disruption in a 5-HT-2B gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the function of the disrupted 5-HT-2B gene in the non-human transgenic animal is modulated.

65. A method of identifying an agent that modulates the expression of 5-HT-2B, the method comprising:

i) providing a cell comprising a disruption in a 5-HT-2B gene;

(b) contacting the cell with an agent; and

(c) determining whether expression of the 5-HT-2B is modulated.

66. A method of identifying an agent that modulates the function of a 5-HT-2B gene, the method comprising:

(a) providing a cell comprising a disruption in a 5-HT-2B gene;

(b) contacting the cell with an agent; and

(c) determining whether the function of the 5-HT-2B gene is modulated.

67. The method of claim 65 or claim 66, wherein the cell is derived from the non-human transgenic animal of claim 60.

68. An agent identified by the method of claim 63, claim 64, claim 65, or claim 66.

69. A transgenic mouse comprising a disruption in a 5-HT-2B gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: embryonic lethality, abnormal embryos, retarded development, and reabsorbed embryos.

70. The transgenic mouse of claim 69, wherein development is arrested at embryonic day 8.5.

71. The transgenic mouse of claim 69, wherein homozygous offspring are undetectable after embryonic day E8.5.

72. The transgenic mouse of claim 69, wherein homozygous embryos die between embryonic day 8.5 and embryonic day 9.5.

73. The transgenic mouse of claim 69, wherein the wherein the embryos are reabsorbed between embryonic day 8.5 and embryonic day 9.5.

74. A method of producing a transgenic mouse comprising a disruption in a 5-HT-2B gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: embryonic lethality, abnormal embryos, retarded development, and reabsorbed embryos, the method comprising:

    (a) introducing a 5-HT-2B gene targeting construct into a cell;

    (b) introducing the cell into a blastocyst;

    (c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

    (d) breeding the chimeric mouse to produce the transgenic mouse comprising a disruption in a 5-HT-2B gene.

75. A transgenic mouse produced by the method of claim 74.

76. A cell derived from the transgenic mouse of claim 60 or claim 75.

77. A method of identifying an agent that ameliorates a phenotype associated with a disruption in a 5-HT-2B gene, the method comprising:

    (a) administering an agent to a transgenic mouse comprising a disruption in a 5-HT-2B gene; and

    (b) determining whether the agent ameliorates at least one of the following phenotypes: embryonic lethality, abnormal embryos, retarded development, and reabsorbed embryos.

78. An agonist or antagonist of a 5-HT-2B receptor.

79. A targeting construct comprising:

    (a) a first polynucleotide sequence homologous to a chemokine receptor 9A gene;

    (b) a second polynucleotide sequence homologous to the chemokine receptor 9A gene; and

    (c) a selectable marker.

80. The targeting construct of claim 79, wherein the targeting construct further comprises a screening marker.

81. A method of producing a targeting construct, the method comprising:

    (a) providing a first polynucleotide sequence homologous to a chemokine receptor 9A gene;

    (b) providing a second polynucleotide sequence homologous to the chemokine receptor 9A;

    (c) providing a selectable marker; and

    (d) inserting the first sequence, second sequence, and selectable marker into a vector, to produce the targeting construct.

82. A method of producing a targeting construct, the method comprising:

    (a) providing a polynucleotide comprising a first sequence homologous to a first region of a chemokine receptor 9A gene and a second sequence homologous to a second region of a chemokine receptor 9A gene;

(b) inserting a positive selection marker in between the first and second sequences to form the targeting construct.

83. A cell comprising a disruption in a chemokine receptor 9A gene.
84. The cell of claim 83, wherein the cell is a murine cell.
85. The cell of claim 84, wherein the murine cell is an embryonic stem cell.
86. A non-human transgenic animal comprising a disruption in a chemokine receptor 9A gene.
87. A cell derived from the non-human transgenic animal of claim 86.
88. A method of producing a transgenic mouse comprising a disruption in a chemokine receptor 9A gene, the method comprising:

    (a) introducing the targeting construct of claim 79 into a cell;

    (b) introducing the cell into a blastocyst;

    (c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

    (d) breeding the chimeric mouse to produce the transgenic mouse.

89. A method of identifying an agent that modulates the expression of a chemokine receptor 9A, the method comprising:

    (a) providing a non-human transgenic animal comprising a disruption in a chemokine receptor 9A gene;

    (b) administering an agent to the non-human transgenic animal; and

    (c) determining whether the expression of chemokine receptor 9A in the non-human transgenic animal is modulated.

90. A method of identifying an agent that modulates the function of a chemokine receptor 9A, the method comprising:

    (a) providing a non-human transgenic animal comprising a disruption in a chemokine receptor 9A gene;

    (b) administering an agent to the non-human transgenic animal; and

    (c) determining whether the function of the disrupted chemokine receptor 9A gene in the non-human transgenic animal is modulated.

91. A method of identifying an agent that modulates the expression of chemokine receptor 9A, the method comprising:

    (a) providing a cell comprising a disruption in a chemokine receptor 9A gene;

    (b) contacting the cell with an agent; and

    (c) determining whether expression of the chemokine receptor 9A is modulated.

92. A method of identifying an agent that modulates the function of a chemokine receptor 9A gene, the method comprising:
    (a) providing a cell comprising a disruption in a chemokine receptor 9A gene;
    (b) contacting the cell with an agent; and
    (c) determining whether the function of the chemokine receptor 9A gene is modulated.
93. The method of claim 91 or claim 92, wherein the cell is derived from the non-human transgenic animal of claim 86.
94. An agent identified by the method of claim 89, claim 90, claim 91, or claim 92.
95. A transgenic mouse comprising a disruption in a chemokine receptor 9A gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: decreased agility, coordination, or balance relative to a wild-type mouse.
96. The transgenic mouse of claim 95, wherein decreased agility, coordination, or balance is characterized by decreased performance on an accelerating rotarod.
97. The transgenic mouse of claim 95, wherein decreased agility, coordination, or balance is characterized by falling from an accelerating rotarod at lower speeds relative to a wild-type mouse.
98. A method of producing a transgenic mouse comprising a disruption in a chemokine receptor 9A gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: decreased agility, coordination, or balance relative to a wild-type mouse, the method comprising:
    (a) introducing a chemokine receptor 9A gene targeting construct into a cell;
    (b) introducing the cell into a blastocyst;
    (c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and
    (d) breeding the chimeric mouse to produce the transgenic mouse comprising a disruption in a chemokine receptor 9A gene.
99. A transgenic mouse produced by the method of claim 98.
100. A cell derived from the transgenic mouse of claim 95 or claim 98.
101. A method of identifying an agent that ameliorates a phenotype associated with a disruption in a chemokine receptor 9A gene, the method comprising:
    (a) administering an agent to a transgenic mouse comprising a disruption in a chemokine receptor 9A gene; and
    (b) determining whether the agent ameliorates at least one of the following phenotypes: decreased agility, coordination, or balance relative to a wild-type mouse.

102. A method of identifying an agent that modulates chemokine receptor 9A expression, the method comprising:

    (a) administering an agent to the transgenic mouse comprising a disruption in a chemokine receptor 9A gene; and

    (b) determining whether the agent modulates chemokine receptor 9A expression in the transgenic mouse, wherein the agent has an effect on at least one of the following behaviors: decreased agility, coordination, or balance relative to a wild-type mouse.

103. A method of identifying an agent that modulates a behavior associated with a disruption in a chemokine receptor 9A gene, the method comprising:

    (a) administering an agent to a transgenic mouse comprising a disruption in a chemokine receptor 9A gene; and

    (b) determining whether the agent modulates agility, coordination, or balance of the transgenic mouse.

104. A method of identifying an agent that modulates chemokine receptor 9A gene function, the method comprising:

    (a) providing a cell comprising a disruption in a chemokine receptor 9A gene;

    (b) contacting the cell with an agent; and

    (c) determining whether the agent modulates chemokine receptor 9A gene function, wherein the agent modulates a phenotype associated with a disruption in a chemokine receptor 9A gene.

105. The method of claim 104, wherein the phenotype comprises at least one of the following: decreased agility, coordination, or balance relative to a wild-type mouse.

106. An agent identified by the method of claim 101, claim 102, claim 103, or claim 104.

107. An agonist or antagonist of a chemokine receptor 9A receptor.

108. Phenotypic data associated with the transgenic mouse of claim 95 or claim 98, wherein the data is in a database.

109. A targeting construct comprising:

    (a) a first polynucleotide sequence homologous to a glucocorticoid-induced receptor gene;

    (b) a second polynucleotide sequence homologous to the glucocorticoid-induced receptor gene; and

    (c) a selectable marker.

110. The targeting construct of claim 111, wherein the targeting construct further comprises a screening marker.

111. A method of producing a targeting construct, the method comprising:

(a) providing a first polynucleotide sequence homologous to a glucocorticoid-induced receptor gene;

(b) providing a second polynucleotide sequence homologous to the glucocorticoid-induced receptor;

(c) providing a selectable marker; and

(d) inserting the first sequence, second sequence, and selectable marker into a vector, to produce the targeting construct.

112. A method of producing a targeting construct, the method comprising:

(a) providing a polynucleotide comprising a first sequence homologous to a first region of a glucocorticoid-induced receptor gene and a second sequence homologous to a second region of a glucocorticoid-induced receptor gene;

(b) inserting a positive selection marker in between the first and second sequences to form the targeting construct.

113. A cell comprising a disruption in a glucocorticoid-induced receptor gene.

114. The cell of claim 113, wherein the cell is a murine cell.

115. The cell of claim 114, wherein the murine cell is an embryonic stem cell.

116. A non-human transgenic animal comprising a disruption in a glucocorticoid-induced receptor gene.

117. A cell derived from the non-human transgenic animal of claim 116.

118. A method of producing a transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene, the method comprising:

(a) introducing the targeting construct of claim 111 into a cell;

(b) introducing the cell into a blastocyst;

(c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

(d) breeding the chimeric mouse to produce the transgenic mouse.

119. A method of identifying an agent that modulates the expression of a glucocorticoid-induced receptor, the method comprising:

(a) providing a non-human transgenic animal comprising a disruption in a glucocorticoid-induced receptor gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the expression of glucocorticoid-induced receptor in the non-human transgenic animal is modulated.

120. A method of identifying an agent that modulates the function of a glucocorticoid-induced receptor, the method comprising:

(a) providing a non-human transgenic animal comprising a disruption in a glucocorticoid-induced receptor gene;

(b) administering an agent to the non-human transgenic animal; and

(c) determining whether the function of the disrupted glucocorticoid-induced receptor gene in the non-human transgenic animal is modulated.

121. A method of identifying an agent that modulates the expression of glucocorticoid-induced receptor, the method comprising:

(a) providing a cell comprising a disruption in a glucocorticoid-induced receptor gene;

(b) contacting the cell with an agent; and

(c) determining whether expression of the glucocorticoid-induced receptor is modulated.

122. A method of identifying an agent that modulates the function of a glucocorticoid-induced receptor gene, the method comprising:

(a) providing a cell comprising a disruption in a glucocorticoid-induced receptor gene;

(b) contacting the cell with an agent; and

(c) determining whether the function of the glucocorticoid-induced receptor gene is modulated.

123. The method of claim 121 or claim 122, wherein the cell is derived from the non-human transgenic animal of claim 116.

124. An agent identified by the method of claim 119, claim 120, claim 121 or claim 122.

125. A transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: hyperactivity, reduced anxiety, decreased propensity toward behavioral despair, or decreased propensity toward depression.

126. The transgenic mouse of claim 125, wherein hyperactivity is characterized by an increase in total distance traveled in an open field test, relative to a wild-type mouse.

127. The transgenic mouse of claim 125, wherein hyperactivity is characterized by an increase in the percent of time spent in the central region of the test chamber in an open field test, relative to a wild-type mouse.

128. The transgenic mouse of claim 125, wherein reduced anxiety is characterized by an increase in total distance traveled in an open field test, relative to a wild-type mouse.

129. The transgenic mouse of claim 125, wherein reduced anxiety is characterized by an increase in the percent of time spent in the central region of the test chamber in an open field test, relative to a wild-type mouse.

130. The transgenic mouse of claim 125, wherein the decreased propensity toward behavioral despair is characterized by less time immobile in a tail suspension test relative to a wild-type mouse.

131. The transgenic mouse of claim 125, decreased propensity toward depression is characterized by less time immobile in a tail suspension test relative to a wild-type mouse.

132. A method of producing a transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene, wherein the transgenic mouse exhibits at least one of the following phenotypes: hyperactivity, reduced anxiety, a decreased propensity toward behavioral despair, or a decreased propensity toward depression, the method comprising:

(a) introducing a glucocorticoid-induced receptor gene targeting construct into a cell;

(b) introducing the cell into a blastocyst;

(c) implanting the resulting blastocyst into a pseudopregnant mouse, wherein said pseudopregnant mouse gives birth to a chimeric mouse; and

(d) breeding the chimeric mouse to produce the transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene.

133. A transgenic mouse produced by the method of claim 132.

134. A cell derived from the transgenic mouse of claim 125 or claim 133.

135. A method of identifying an agent that ameliorates a phenotype associated with a disruption in a glucocorticoid-induced receptor gene, the method comprising:

(a) administering an agent to a transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene; and

(b) determining whether the agent ameliorates at least one of the following phenotypes: hyperactivity, reduced anxiety, decreased propensity toward behavioral despair, or decreased propensity toward depression.

136. A method of identifying an agent that modulates glucocorticoid-induced receptor expression, the method comprising:

(a) administering an agent to the transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene; and

(b) determining whether the agent modulates glucocorticoid-induced receptor expression in the transgenic mouse, wherein the agent has an effect on at least one of the following behaviors: hyperactivity, anxiety, behavioral despair, or depression.

137. A method of identifying an agent that modulates a behavior associated with a disruption in a glucocorticoid-induced receptor gene, the method comprising:

    (a) administering an agent to a transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene; and

    (b) determining whether the agent modulates hyperactivity, anxiety, behavioral despair, or depression.

138. A method of identifying an agent that modulates glucocorticoid-induced receptor gene function, the method comprising:

    (a) providing a cell comprising a disruption in a glucocorticoid-induced receptor gene;

    (b) contacting the cell with an agent; and

    (c) determining whether the agent modulates glucocorticoid-induced receptor gene function, wherein the agent modulates a phenotype associated with a disruption in a glucocorticoid-induced receptor gene.

139. The method of claim 138, wherein the phenotype comprises at least one of the following: hyperactivity, reduced anxiety, decreased propensity toward behavioral despair, or decreased propensity toward depression.

140. An agent identified by the method of claim 135, claim 136, claim 137, or claim 138.

141. A transgenic mouse comprising a disruption in a glucocorticoid-induced receptor gene, wherein the transgenic mouse exhibits hyperactivity, reduced anxiety, a decreased propensity toward behavioral despair, or a decreased propensity toward depression relative to a wild-type mouse.

142. An agonist or antagonist of a glucocorticoid-induced receptor.

143. Phenotypic data associated with the transgenic mouse of claim 125 or claim 133, wherein the data is in a database.

TCTAGACTGGACAGCATCCACAAGAGAAGCACCTAGAAGGAGAATTTTCCCCAGCAGCTTGCTCAGGACCCTGCA
GGAGCCGCAGCTGGGACTGGACCTGCTGTTAACCATGAACTCTTCCTGCTGCCTGTCTTCTGTTTCTCCGATGCT
GCCTAACCTCTCTGAGCACCCTGCAGCCCCTCCTGCCAGCAACCGGAGCGGCAGTGGGTTCTGTGAGCAGGTCTT
CATCAAGCCGGAGGTCTTCCTGGCTCTGGGCATCGTCAGTCTGATGGAAAACATCCTGGTGATCCTGGCTGTGGT
CAGGAATGGCAACCTGCACTCTCCCATGTACTTCTTCCTGTGCAGCCTGGCTGCAGCCGACATGCTGGTGAGCCT
GTCCAACTCCCTGGAGACCATCATGATCGCCGTGATCAACAGCGACTCCCTGACCTTGGAGGACCAGTTTATCCA
GCACATGGATAATATCTTCGACTCTATGATTTGCATCTCCCTGGTGGCCTCCATCTGCAACCTCCTGGCCATTGC
CATCGACAGGTACGTCACCATCTTCTATGCCCTTCGGTACCACAGCATCATGACAGTTAGGAAAGCCCTCACCTT
GATCGGGGTCATCTGGGTCTGCTGCGGCATCTGCGGCGTGATGTTCATCATCTACTCCGAGAGCAAGATGGTCAT
CGTGTGTCTCATCACCATGTTCTTCGCCATGGTGCTCCTCATGGGCACCCTATATATCCACATGTTCCTCTTCGC
CAGGCTCCACGTCCAGCGCATCGCAGTGCTGCCCCCTGCTGGCGTGGTGGCCCCACAGCAGCACTCCTGCATGAA
GGGGGCTGTCACCATCACTATCCTGCTGGGTGTTTTCATCTTCTGCTGGGCGCCTTTCTTCCTCCACCTGGTCCT
CATCATCACCTGCCCCACCAATCCCTACTGCATCTGCTACACGGCCCATTTCAACACCTACCTGGTTCTCATCAT
GTGCAACTCCGTCATCGACCCCCTCATCTACGCCTTCCGCAGCCTGGAGCTGCGCAACACGTTCAAGGAGATTCT
CTGCGGCTGCAACAGCATGAACTTGGGCTAGGATGCCCGTGGAGGTGTTCCACATCCAGCCAAGAGACAAAAACA
ACGCTCAGACGGGACGTAAAAGGGTGTTAGGAGCTGGAACTGTGCTTGGCTTCGTCTGTAAGCTCGTGGCCCTTT
GCAGACGGGACACGGCGTAGGATGGGCTGTCTGTGAGGATCTGTGTGTGGGTAAGTCAGTTTGATCTAGCACATA
GCCTGGAAGAATCAGGCAAAGCAGCCCTGAGTGTCATCTGTGTTCATTGCTAGGCACCCAGGGTTTGTGGCCCCT
GCCTGCTTATTGGCTTTGTACCAGTAACTGTGCTTCAAGCCAACCAGACCGGAGGGCTCTCGTGAGCAGAAAGAG
TGCTTAGACTTCCGGCAAGCATCCTGGCTCACAGCGGCCACCTCCTGACCACTACCGGGAGAGCTTTGCACATAT
TCTGTGGGAGATTGAGTGAAGCCCTGAAAACAATGTGATATTTGCTGCTCCCTTCCAGAACTTACATCTGTGCCA
GCCTCCCCGAACCCCTGCACAGAGACATGACCCCCTTCTCCCTGTGCCGTTGTCATGGTTGTTATTATTGTTGGA
GTTTTGTTCGTTAAAATCTAAGCTT (SEQ ID NO:1)

MNSSCCLSSVSPMLPNLSEHPAAPPASNRSGSGFCEQVFIKPEVFLALGIVSLMENILVILAVVRNGNLHSPMYF
FLCSLAAADMLVSLSNSLETIMIAVINSDSLTLEDQFIQHMDNIFDSMICISLVASICNLLAIAIDRYVTIFYAL
RYHSIMTVRKALTLIGVIWVCCGICGVMFIIYSESKMVIVCLITMFFAMVLLMGTLYIHMFLFARLHVQRIAVLP
PAGVVAPQQHSCMKGAVTITILLGVFIFCWAPFFLHLVLIITCPTNPYCICYTAHFNTYLVLIMCNSVIDPLIYA
FRSLELRNTFKEILCGCNSMNLG (SEQ ID NO:2)

FIGURE 1

Underlined = deleted in targeting construct

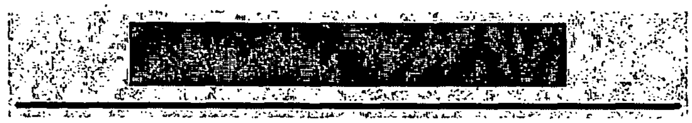
Bold = sequence flanking Neo insert in targeting construct

TCTAGACTGGACAGCATCCACAAGAGAAGCACCTAGAAGGAGAATTTTCCCCAGCAGCTT
GCTCAGGACCCTGCAGGAGCCGCAGCTGGGACTGGACCTGCTGTTAACCATGAACTCTTC
CTGCTGCCTGTCTTCTGTTTCTCCGATGCTGCCTAACCTCTCTGAGCACCCTGCAGCCCC
TCCTGCCAGCAACCGGAGCGGCAGTGGGTTCTGTGAGCAGGTCTTCATCAAGCCGGAGGT
CTTCCTGGCTCTGGGCATCGTCAGTCTGATGGAAAACATCCTGGTGATCCTGGCTGTGGT
CAGGAATGGCAACCTGCACTCTCCCATGTACTTCTTCCTGTGCAGCCTGGCTGCAGCCGA
CATGCTGGTGAGCCTGTCCAACTCCCTGGAGACCATCATGATCGCCGTGATCAACAGCGA
CTCCCTGACCTTGGAGGACCAGTTTATCCAGCACATGGATAATATCTTCGACTCTATGAT
TTGCATCTCCCTGGTGGCCTCCATCTGCAACCTCCTGGC**CATTGCCATCGACAGGTACGT**
CACCATCTTCTATGCCCTTCGGTACCACAGCATCATGACAGTTAGGAAAGCCCTCACCTT
GATCGGGGTCATCTGGGTCTGCTGCGGCATCTGCGGCGTGATGTTCATCATCTACTCCGA
GAGCAAGATGGTCATCGTGTGTCTCATCACCATGTTCTTCGCCATGGTGCTCCTCATGGG
CACCCTATATATCCACATGTTCCTCTTCGCCAGGCTCCACGTCCAGCGCATCGCAGTGCT
GCCCCCTGCTGGCGTGGTGGCCCCACAGCAGCACTCCTGCATGAAGGGGGCTGTCACCAT
CACTATCCTGCTGGGTGTTTTCATCTTCTGCTGGGCGCCTTTCTTCCTCCACCTGGTCCT
CATCATCACCTGCCCCACCAATCCCTACTGCATCTGCTACACGGCCCATTTCAACACCTA
CCTGGTTCTCATCATGTGCAACTCCGTCATCGACCCCCTCATCTACGCCTTCCGCAGCCT
GGAGCTGCGCAACACGTTCAAGGAGATTCTCTGCGGCTGCAACAGCATGAACTTGGGCTA
GGATGCCCGTGGAGGTGTTCCACATCCAGCCAAGAGACAAAAACAACGCTCAGACGGGAC
GTAAAAGGGTGTTAGGAGCTGGAACTGTGCTTGGCTTCGTCTGTAAGCTCGTGGCCCTTT
GCAGACGGGACACGGCGTAGGATGGGCTGTCTGTGAGGATCTGTGTGTGGGTAAGTCAGT
TTGATCTAGCACATAGCCTGGAAGAATCAGGCAAAGCAGCCCTGAGTGTCATCTGTGTTC
ATTGCTAGGCACCCAGGGTTTGTGGCCCCTGCCTGCTTATTGGCTTTGTACCAGTAACTG
TGCTTCAAGCCAACCAGACCGGAGGGCTCTCGTGAGCAGAAAGAGTGCTTAGACTTCCGG
CAAGCATCCTGGCTCACAGCGGCCACCTCCTGACCACTACCGGGAGAGCTTTGCACATAT
TCTGTGGGAGATTGAGTGAAGCCCTGAAAACAATGTGATATTTGCTGCTCCCTTCCAGAA
CTTACATCTGTGCCAGCCTCCCCGAACCCCTGCACAGAGACATGACCCCCTTCTCCCTGT
GCCGTTGTCATGGTTGTTATTATTGTTGGAGTTTTGTTCGTTAAAATCTAAGCTT

FIGURE 2A

Gene Sequence Structure * 296 bp Sequence Deleted 519 bp

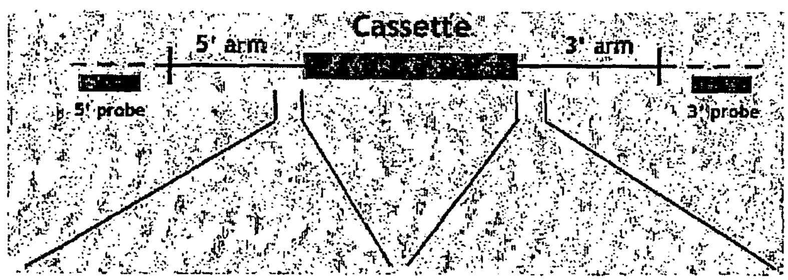
Size of CDS: 1675 bp

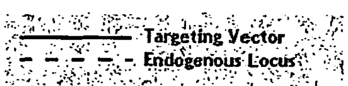


Targeting Vector*
(genomic sequence)

Construct Number: 2787

Arm Length:
5': 1.6 kb
3': 4 kb





* Not drawn to scale

5'>ACCTGCTGTTAACCATGAACT CTTCCTGCTGCCTGTCTTCTGTTT CTCCGATGCTGCCTAACCTCTCTG AGCACCCTGCAGCCCCTCCTGCCA GCAACCGGAGCGGCAGTGGGTTCT GTGAGCAGGTCTTCATCAAGCCGG AGGTCTTCCTGGCTCTGGGCATCG TCAGTCTGATGGAAAACATCCTGG TGATCCTGGCT<3' (SEQ ID NO:3)	5'>CATTGCCATCGACAGGTACGT CACCATCTTCTATGCCCTTCGGTA CCACAGCATCATGACAGTGAGGAA AGCCCTCACCTTGATCGGGGTCAT CTGGGTCTGCTGCGGCATCTGCGG CGTGATGTTCATCATCTACTCCGA GAGCAAGATGGTCATCGTGTGTCT CATCACCATGTTCTTCGCCATGGT GCTCCTCATGG<3' (SEQ ID NO:4)

FIGURE 2B

ACTGTCTGGAACTGGACTGAGTCACCAAAAGGCGAATGGCTTCATCTTATAAAATGTCTGAACAAAGCACAACTT
CTGAGCACATTTTACAGAAGACATGTGATCACCTGATCCTGACTAACCGTTCTGGATTAGAGACAGACTCAGTAG
CAGAGGAAATGAAGCAGACTGTGGAGGGACAGGGGCATACAGTGCACTGGGCAGCTCTCCTGATACTCGCGGTGA
TAATACCCACCATTGGTGGGAACATCCTTGTGATTCTGGCTGTTGCACTGGAGAAAAGGCTGCAGTACGCTACCA
ACTACTTTTTAATGTCCTTGGCGATAGCAGATTTGCTGGTTGGATTGTTTGTGATGCCGATTGCCCTCTTGACAA
TCATGTTTGAGGCTATATGGCCCCTCCCACTGGCCCTGTGTCCTGCCTGGTTATTCCTCGATGTTCTCTTTTCAA
CTGCCTCCATCATGCATCTCTGTGCCATTTCCCTGGACCGCTATATAGCCATCAAAAAGCCAATTCAGGCCAATC
AGTGCAACACCCGGGCTACTGCATTCATCAAGATTACAGTGGTATGGTTAATTTCAATAGGCATCGCCATCCCAG
TCCCTATTAAAGGAATCGAGACTGATGTGATTAATCCACACAATGTCACCTGTGAGCTGACAAAGGACCGCTTTG
GCAGTTTTATGGTCTTTGGGTCACTGGCTGCTTTCTTCGTACCTCTCACCATCATGGTAGTCACTTACTTTCTCA
CCATTCACACTTTACAGAAGAAAGCTTACTTGGTCAAAAATAAGCCACCTCAACGCCTAACACGGTGGACTGTGC
CCACAGTTTTCCTAAGGGAAGACTCATCCTTTTCATCACCAGAAAAGGTGGCAATGCTGGATGGGTCTCACAGGG
ATAAAATTCTACCTAACTCAAGTGATGAGACACTTATGCGAAGAATGTCCTCAGTTGGAAAAAGATCAGCCCAAA
CCATTTCTAATGAGCAGAGAGCCTCGAAGGCCCTTGGAGTCGTGTTTTTCCTTTTTCTGCTTATGTGGTGCCCCT
TTTTTATTACAAATCTAACTTTAGCTCTGTGTGATTCCTGCAATCAGACCACTCTCAAAACACTCCTGGAGATAT
TTGTGTGGATAGGCTACGTTTCCTCGGGGGTGAATCCTCTGATCTATACACTCTTCAATAAGACATTTCGGGAAG
CATTTGGCAGGTACATCACCTGCAATTACCGAGCCACAAAGTCAGTAAAAGCACTTAGGAAGTTTTCCAGTACAC
TTTGTTTTGGGAATTCAATGGTAGAAAACTCTAAATTTTTCACAAAACATGGAATTCGAAATGGGATCAACCCTG
CCATGTACCAGAGCCCAATGAGGCTCCGATGTTCAACCATTCAGTCCTCATCAATCATCCTCCTCGATACCCTTC
TCACTGAAAACGATGGCGACAAAGCGGAAGAGCAGGTCAGCTACATATTGCAGGAACGGGCCGGCCTCATCTTGA
GAGAGGGTGATGAGCAGGACGCACGCGCACCATGGCAGGTTCAAGAGTGA (SEQ ID NO:5)

MASSYKMSEQSTTSEHILQKTCDHLILTNRSGLETDSVAEEMKQTVEGQGHTVHWAALLILAVIIPTIGGNILVI
LAVALEKRLQYATNYFLMSLAIADLLVGLFVMPIALLTIMFEAIWPLPLALCPAWLFLDVLFSTASIMHLCAISL
DRYIAIKKPIQANQCNTRATAFIKITVVWLISIGIAIPVPIKGIETDVINPHNVTCELTKDRFGSFMVFGSLAAF
FVPLTIMVVTYFLTIHTLQKKAYLVKNKPPQRLTRWTVPTVFLREDSSFSSPEKVAMLDGSHRDKILPNSSDETL
MRRMSSVGKRSAQTISNEQRASKALGVVFFLFLLMWCPFFITNLTLALCDSCNQTTLKTLLEIFVWIGYVSSGVN
PLIYTLFNKTFREAFGRYITCNYRATKSVKALRKFSSTLCFGNSMVENSKFFTKHGIRNGINPAMYQSPMRLRCS
TIQSSSIILLDTLLTENDGDKAEEQVSYILQERAGLILREGDEQDARAPWQVQE (SEQ ID NO:6)

FIGURE 3

Underlined = deleted in targeting construct

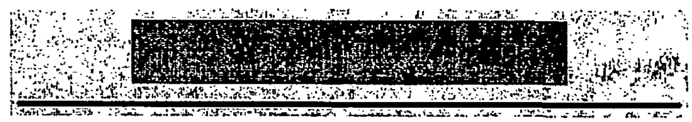
Bold = sequence flanking Neo insert in targeting construct

ACTGTCTGGAACTGGACTGAGTCACCAAAAGGCGAATGGCTTCATCTTATAAAATGTCTG
AACAAAGCACAACTTCTGAGCACATTTTACAGAAGACATGTGATCACCTGATCCTGACTA
ACCGTTCTGGATTAGAGACAGACTCAGTAGCAGAGGAAATGAAGCAGACTGTGGAGGGAC
AGGGGCATACAGTGCACTGGGCAGCTCTCCTGATACTCGCGGTGATAATACCCACCATTG
GTGGGAACATCCTTGTGATTCTGGCTGTTGCACTGGAGAAAAGGCTGCAGTACGCTACCA
ACTACTTTTTAATGTCCTT**GGCGATAGCAGATTTGCTGGTTGGATTGTTTGTGATGCCGA**
TTGCCCTCTTGACAATCATGTTTGAGGCTATATGGCCCCTCCCACTGGCCCTGTGTCCTG
CCTGGTTATTCCTCGATGTTCTCTTTTCAACTGCCTCCATCATGCATCTCTGTGCCATTT
CCCTGGACCGCTATATAGCCATCAAAAAGCCAATTCAGGCCAATCAGTGCAACACCCGGG
CTACTGCATTCATCAAGATTACAGTGGTATGGTTAATTTCAATAGGCATCGCCATCCCAG
TCCCTATTAAAGGAATCGAGACTGATGTGATTAATCCACACAATGTCACCTGTGAGCTGA
CAAAGGACCGCTTTGGCAGTTTTATGGTCTTTGGGTCACTGGCTGCTTTCTTCGTACCTC
TCACCATCATGGTAGTCACTTACTTTCTCACCATTCACACTTTACAGAAGAAAGCTTACT
TGGTCAAAAATAAGCCACCTCAACGCCTAACACGGTGGACTGTGCCCACAGTTTTCCTAA
GGGAAGACTCATCCTTTTCATCACCAGAAAAGGTGGCAATGCTGGATGGGTCTCACAGGG
ATAAAATTCTACCTAACTCAAGTGATGAGACACTTATGCGAAGAATGTCCTCAGTTGGAA
AAAGATCAGCCCAAACCATTTCTAATGAGCAGAGAGCCTCGAAGGCCCTTGGAGTCGTGT
TTTTCCTTTTTCTGCTTATGTGGTGCCCCTTTTTTATTACAAATCTAACTTTAGCTCTGT
GTGATTCCTGCAATCAGACCACTCTCAAAACACTCCTGGAGATATTTGTGTGGATAGGCT
ACGTTTCCTCGGGGGTGAATCCTCTGATCTATACACTCTTCAATAAGACATTTCGGAAG
CATTTGGCAGGTACATCACCTGCAATTACCGAGCCACAAAGTCAGTAAAAGCACTTAGGA
AGTTTTCCAGTACACTTTGTTTTGGGAATTCAATGGTAGAAAACTCTAAATTTTTCACAA
AACATGGAATTCGAAATGGGATCAACCCTGCCATGTACCAGAGCCCAATGAGGCTCCGAT
GTTCAACCATTCAGTCCTCATCAATCATCCTCCTCGATACCCTTCTCACTGAAAACGATG
GCGACAAAGCGGAAGAGCAGGTCAGCTACATATTGCAGGAACGGGCCGGCCTCATCTTGA
GAGAGGGTGATGAGCAGGACGCACGCGCACCATGGCAGGTTCAAGAGTGA

FIGURE 4A

Gene Sequence Structure * 130 bp Sequence Deleted 319 bp

Size of full-length cDNA: 1550 bp

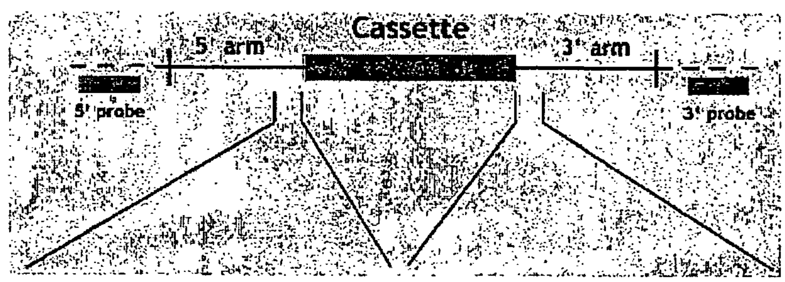


Targeting Vector*
(genomic sequence)

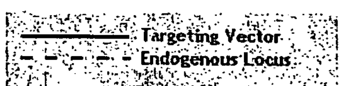
Construct Number: 2520

Arm Length:
5': 1.6 kb
3': 5 kb

LacZ-Neo



5'>TGAGTGTCTGGTGGGTTTGCTAAATGCTTTGCTAAAGCAGATGACTTGCTTAGCTACTGACCATGCTGACCACTGTCTGGAACTGGACTGAGTCACCAAAAGGCGAATGGCTTCATCTTATAAAATGTCTGAACAAAGCACAACTTCTGAGCACATTTTACAGAAGACATGTGATCACCTGATCCTGACTAACCGTTCTG<3' (SEQ ID NO:7)	5'>GGCGATAGCAGATTTGCTGGTTGGATTGTTTGTGATGCCGATTGCCCTCTTGACAATCATGTTTGGTGAGTATTTCCCCTTGTTCCTGCCACTGAACACTACTAACGTAGTGAAATGGACACTCACTGACCTTTATTTTGTTTGAAATAAAAGAAGGACCTGGATTAAAAACACAGAAGGGAACATTCCTTCATTTTTCA<3' (SEQ ID NO:8)



* Not drawn to scale

FIGURE 4B

AATATTTTCCTTGACCTAATGCCATCTTGTGTCCCCTTGCAGAGCCCTATTCCTAACATGGCTGATGACTATGGC
TCTGAATCCACATCTTCCATGGAAGACTACGTTAACTTCAACTTCACTGACTTCTACTGTGAGAAAAACAATGTC
AGGCAGTTTGCGAGCCATTTCCTCCCACCCTTGTACTGGCTCGTGTTCATCGTGGGTGCCTTGGGCAACAGTCTT
GTTATCCTTGTCTACTGGTACTGCACAAGAGTGAAGACCATGACCGACATGTTCCTTTTGAATTTGGCAATTGCT
GACCTCCTCTTTCTTGTCACTCTTCCCTTCTGGGCCATTGCTGCTGCTGACCAGTGGAAGTTCCAGACCTTCATG
TGCAAGGTGGTCAACAGCATGTACAAGATGAACTTCTACAGCTGTGTGTTGCTGATCATGTGCATCAGCGTGGAC
AGGTACATTGCCATTGCCCAGGCCATGAGAGCACATACTTGGAGGGAGAAAAGGCTTTTGTACAGCAAAATGGTT
TGCTTTACCATCTGGGTATTGGCAGCTGCTCTCTGCATCCCAGAAATCTTATACAGCCAAATCAAGGAGGAATCC
GGCATTGCTATCTGCACCATGGTTTACCCTAGCGATGAGAGCACCAAACTGAAGTCAGCTGTCTTGACCCTGAAG
GTCATTCTGGGGTTCTTCCTTCCCTTCGTGGTCATGGCTTGCTGCTATACCATCATCATTCACACCCTGATACAA
GCCAAGAAGTCTTCCAAGCACAAAGCCCTAAAAGTGACCATCACTGTCCTGACCGTCTTTGTCTTGTCTCAGTTT
CCCTACAACTGCATTTTGTTGGTGCAGACCATTGACGCCTATGCCATGTTCATCTCCAACTGTGCCGTTTCCACC
AACATTGACATCTGCTTCCAGGTCACCCAGACCATCGCCTTCTTCCACAGTTGCCTGAACCCTGTTCTCTATGTT
TTTGTGGGTGAGAGATTCCGCCGGGATCTCGTGAAAACCCTGAAGAACTTGGGTTGCATCAGCCAGGCCCAGTGG
GTTTCATTTACAAGGAGAGAGGGAAGCTTGAAGCTGTCGTCTATGTTGCTGGAGACAACCTCAGGAGCACTCTCC
CTCTGAGGGGTCTTCTCTGAGGTGCATGGTTCTTTTGGAAGAAATGAGAAATACATGAAACAGTTTCCCCACTGA
TGGGACCAGAGAGAGTGAAAGAGAAAAGAAAACTCAGAAAGGGATGAATCTGAACTATATGATTACTTGTAGTCA
GAATTTGCCAAAGCAAATATTTCAAAATCAACTGACTAGTGCAGGAGGCTGTTGATTGGCTCTTGACTGTGATGC
CCGCAATTCTCAAAGGAGGACTAAGGACCGGCACTGTGGAGCACCCTGGCTTTGCCACTCGCCGGAGCATCAATG
CCGCTGCCTCTGGAGGAGCCCTTGGATTTTCTCCATGCACTGTGAACTTCTGTGGCTTCAGTTCTCATGCTGCCT
CTTCCAAAAGGGGACACAGAAGCACTGGCTGCTGCTACAGACCGCAAAAGCAGAAAGTTTCGTGAAAATGTCCAT
CTTTGGGAAATTTTCTACCCTGCTCTTGAGCCTGATAACCCATGCCAGGTCTTATAGATTCCTGATCTAGAACCT
TTCCAGGCAATCTCAGACCTAATTTCCTTCTGTTCTCCTTGTTCTGTTCTGGGCCAGTGAAGGTCCTTGTTCTGA
TTTTGAAACGATCTGCAGGTCTTGCCAGTGAACCCCTGGACAACTGACCACACCCACAAGGCATCCAAAGTCTGT
TGGCTTCCAATCCATTTCTGTGTCCTGCTGGAGGTTTTAACCTAGACAAGGATTCCGCTTATTCCTTGGTATGGT
GACAGTGTCTCTCCATGGCCTGAGCAGGGAGATTATAACAGCTGGGTTCGCAGGAGCCAGCCTTGGCCCTGTTGT
AGGCTTGTTCTGTTGAGTGGCACTTGCTTTGGGTCCACCGTCTGTCTGCTCCCTAGAAAATGGGCTGGTTCTTTT
GGCCCTCTTCTTTCTGAGGCCCACTTTATTCTGAGGAATACAGTGAGCAGATATGGGCAGCAGCCAGGTAGGGCA
AAGGGGTGAAGCGCAGGCCTTGCTGGAAGGCTATTTACTTCCATGCTTCTCCTTTTCTTACTCTATAGTGGCAAC
ATTTTAAAAGCTTTTAACTTAGAGATTAGGCTGAAAAAAATAAGTAATGGAATTCACCTTTGCATCTTTTGTGTC
TTTCTTATCATGATTTGGCAAAATGCATCACCTTTGAAAATATTTCACATATTGGAAAAGTGCTTTTTAATGTGT
ATATGAAGCATTAATTACTTGTCACTTTCTTTACCCTGTCTCAATATTTTAAGTGTGTGCAATTAAAGATCAAAT
AGATACATTAAGAGTGTGAAGGCTGGTCTGAAGGTAGTGAGCTATCTCAATCGGATTGTTCACACTCAGTTACAG
ATTGAACTCCTTGTTCTACTTCCCTGCTTCTCTCTACTGCAATTGACTAGTCTTTAAAAAAAAGTGTGAAGAGTA
AGCAATAGGGATAAGGAAATAAGATCT (SEQ ID NO:9)

MADDYGSESTSSMEDYVNFNFTDFYCEKNNVRQFASHFLPPLYWVFIVGALGNSLVILVYWYCTRVKTMTDMFLL
NLAIADLLFLVTLPFWAIAAADQWKFQTFMCKVVNSMYKMNFYSCVLLIMCISVDRYIAIAQAMRAHTWREKRLL
YSKMVCFTIWVLAAALCIPEILYSQIKEESGIAICTMVYPSDESTKLKSAVLTLKVILGFFLPFVVMACCYTIII
HTLIQAKKSSKHKALKVTITVLTVFVLSQFPYNCILLVQTIDAYAMFISNCAVSTNIDICFQVTQTIAFFHSCLN
PVLYVFVGERFRRDLVKTLKNLGCISQAQWVSFTRREGSLKLSSMLLETTSGALSL (SEQ ID NO:10)

FIGURE 5

<u>Underlined</u> = deleted in targeting construct

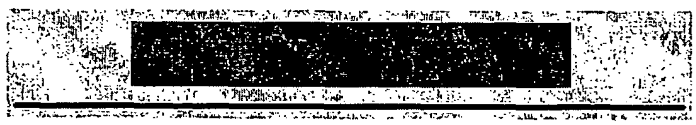
Bold = sequence flanking Neo insert in targeting construct

AATATTTTCCTTGACCTAATGCCATCTTGTGTCCCCTTGCAGAGCCCTATTCCTAACATG
GCTGATGACTATGGCTCTGAATCCACATCTTCCATGGAAGACTACGTTAACTTCAACTTC
ACTGACTTCTACTGTGAGAAAAACAATGTCAGGCAGTTTGCGAGCCATTTCCTCCCACCC
TTGTACTGGCTCGTGTTCATCGTGGGTGCCTTGGGCAACAGTCTTGTTAT<u>CCTTGTCTAC</u>
<u>TGGTACTGCACAAGAGTGAAGACCATGACCGACATGTTCCTTTTGAATTTGGCAATTGCT</u>
<u>GACCTCCTCTTTCTTGTCACTCTTCCCTTCTGGGCCATTGCTGCTGCTGACCAGTGGAAG</u>
<u>TTCCAGACCTTCATGTGCAAGGTGGTCAACAGCA</u>**TGTACAAGATGAACTTCTACAGCTGT**
GTGTTGCTGATCATGTGCATCAGCGTGGACAGGTACATTGCCATTGCCCAGGCCATGAGA
GCACATACTTGGAGGGAGAAAAGGCTTTTGTACAGCAAAATGGTTTGCTTTACCATCTGG
GTATTGGCAGCTGCTCTCTGCATCCCAGAAATCTTATACAGCCAAATCAAGGAGGAATCC
GGCATTGCTATCTGCACCATGGTTTACCCTAGCGATGAGAGCACCAAACTGAAGTCAGCT
GTCTTGACCCTGAAGGTCATTCTGGGGTTCTTCCTTCCCTTCGTGGTCATGGCTTGCTGC
TATACCATCATCATTCACACCCTGATACAAGCCAAGAAGTCTTCCAAGCACAAAGCCCTA
AAAGTGACCATCACTGTCCTGACCGTCTTTGTCTTGTCTCAGTTTCCCTACAACTGCATT
TTGTTGGTGCAGACCATTGACGCCTATGCCATGTTCATCTCCAACTGTGCCGTTTCCACC
AACATTGACATCTGCTTCCAGGTCACCCAGACCATCGCCTTCTTCCACAGTTGCCTGAAC
CCTGTTCTCTATGTTTTTGTGGGTGAGAGATTCCGCCGGGATCTCGTGAAAACCCTGAAG
AACTTGGGTTGCATCAGCCAGGCCCAGTGGGTTTCATTTACAAGGAGAGAGGGAAGCTTG
AAGCTGTCGTCTATGTTGCTGGAGACAACCTCAGGAGCACTCTCCCTCTGAGGGGTCTTC
TCTGAGGTGCATGGTTCTTTTGGAAGAAATGAGAAATACATGAAACAGTTTCCCCACTGA
TGGGACCAGAGAGAGTGAAAGAGAAAAGAAAACTCAGAAAGGGATGAATCTGAACTATAT
GATTACTTGTAGTCAGAATTTGCCAAAGCAAATATTTCAAAATCAACTGACTAGTGCAGG
AGGCTGTTGATTGGCTCTTGACTGTGATGCCCGCAATTCTCAAAGGAGGACTAAGGACCG
GCACTGTGGAGCACCCTGGCTTTGCCACTCGCCGGAGCATCAATGCCGCTGCCTCTGGAG
GAGCCCTTGGATTTTCTCCATGCACTGTGAACTTCTGTGGCTTCAGTTCTCATGCTGCCT
CTTCCAAAAGGGGACACAGAAGCACTGGCTGCTGCTACAGACCGCAAAAGCAGAAAGTTT
CGTGAAAATGTCCATCTTTGGGAAATTTTCTACCCTGCTCTTGAGCCTGATAACCCATGC
CAGGTCTTATAGATTCCTGATCTAGAACCTTTCCAGGCAATCTCAGACCTAATTTCCTTC
TGTTCTCCTTGTTCTGTTCTGGGCCAGTGAAGGTCCTTGTTCTGATTTTGAAACGATCTG
CAGGTCTTGCCAGTGAACCCCTGGACAACTGACCACACCCACAAGGCATCCAAAGTCTGT
TGGCTTCCAATCCATTTCTGTGTCCTGCTGGAGGTTTTAACCTAGACAAGGATTCCGCTT
ATTCCTTGGTATGGTGACAGTGTCTCTCCATGGCCTGAGCAGGGAGATTATAACAGCTGG
GTTCGCAGGAGCCAGCCTTGGCCCTGTTGTAGGCTTGTTCTGTTGAGTGGCACTTGCTTT
GGGTCCACCGTCTGTCTGCTCCCTAGAAAATGGGCTGGTTCTTTTGGCCCTCTTCTTTCT
GAGGCCCACTTTATTCTGAGGAATACAGTGAGCAGATATGGGCAGCAGCCAGGTAGGGCA
AAGGGGTGAAGCGCAGGCCTTGCTGGAAGGCTATTTACTTCCATGCTTCTCCTTTTCTTA
CTCTATAGTGGCAACATTTTAAAAGCTTTTAACTTAGAGATTAGGCTGAAAAAAATAAGT
AATGGAATTCACCTTTGCATCTTTTGTGTCTTTCTTATCATGATTTGGCAAAATGCATCA
CCTTTGAAAATATTTCACATATTGGAAAAGTGCTTTTTAATGTGTATATGAAGCATTAAT
TACTTGTCACTTTCTTTACCCTGTCTCAATATTTTAAGTGTGTGCAATTAAAGATCAAAT
AGATACATTAAGAGTGTGAAGGCTGGTCTGAAGGTAGTGAGCTATCTCAATCGGATTGTT
CACACTCAGTTACAGATTGAACTCCTTGTTCTACTTCCCTGCTTCTCTCTACTGCAATTG
ACTAGTCTTTAAAAAAAAGTGTGAAGAGTAAGCAATAGGGATAAGGAAATAAGATCT

FIGURE 6A

Gene Sequence Structure * 231 bp Sequence Deleted 394 bp

Size of CDS: 2577 bp

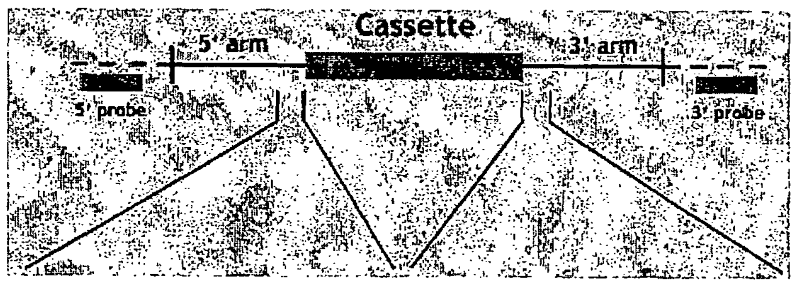


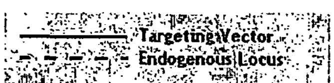
Targeting Vector*
(genomic sequence)

Construct Number: 3340

Arm Length:
5': 2.7 kb
3': 4 kb

LacZ-Neo





* Not drawn to scale

5'>TATTCCTTACAGAGCCTTATT
CCTGGCATGTTTGATGACTTCAGC
TATGACTCCACTGCTTCCACAGAT
GACTACATGAATTTGAATTTCAGT
AGCTTCTTCTGTAAGAAAAATAAT
GTCAGGCAGTTTGCAAGCCATTTT
CTCCCACCTCTGTACTGGCTTGTG
TTCATTGTGGGCACCTTGGGCAAC
AGCCTGGTCAT<3'
(SEQ.ID.NO.11)

5'>TGTACAAGATGAACTTCTACA
GCTGTGTGCTTCTCATCATGTGCA
TCAGTGTGGACAGATACATTGCCA
TTGTACAGGCCATGAAGGCTCAGG
TCTGGAGGCAGAAAAGGCCGCTAT
ACAGCAAGATGGTCTGCATTACCA
TCTGGGTGATGGCAGCTGTGCTCT
GCACCCCAGAAATCCTGTACAGTC
AAGTCAGTGGG<3'
(SEQ.ID.NO.12)

FIGURE 6B

CATGAAGGTTCCTCCTGTCCTGCTTCTCTTTCTTCTGTCCTCAGTGCGAGCTACTGAGCAACCGCAGGTCGTCAC
TGAGCATCCCAGCATGGAGGCAGCCCTGACCGGGCCCAACGCCTCCTCGCACTTCTGGGCCAACTACACTTTCTC
TGACTGGCAGAACTTCGTGGGCAGGAGACGTTATGGGGCCGAGTCCCAGAACCCCACGGTGAAAGCACTGCTCAT
CGTGGCCTACTCATTCACCATCGTCTTCTCGCTCTTCGGTAATGTCCTGGTCTGTCATGTCATCTTCAAGAACCA
GCGCATGCACTCGGCCACCAGCCTCTTCATTGTCAACCTGGCAGTGGCGGACATCATGATCACATTGCTCAACAC
GCCCTTCACTTTGGTCCGCTTTGTGAACAGCACATGGGTGTTTGGGAAGGGCATGTGTCATGTCAGTCGCTTTGC
TCAGTACTGTTCTCTACATGTCTCAGCACTGACTCTGACAGCTATCGCAGTGGACCGCCACCAGGTCATCATGCA
TCCACTGAAGCCTCGGATCTCCATCACCAAGGGTGTCATATATATTGCTGTCATCTGGGTCATGGCTACCTTCTT
CTCTCTGCCACATGCCATCTGCCAGAAACTGTTTACCTTCAAGTACAGTGAGGACATTGTGCGCTCCCTCTGCCT
GCCGGACTTCCCGGAGCCAGCTGACCTCTTCTGGAAGTATCTGGACCTGGCCACCTTCATCCTGCTCTACCTACT
TCCACTCTTCATTATCTCAGTGGCCTATGCTCGTGTGGCCAAGAAGCTGTGGCTCTGTAACACCATTGGCGACGT
GACCACAGAGCAGTACCTCGCCCTGCGACGCAAGAAGAAGACCACCGTGAAGATGCTGGTGCTTGTGGTAGTCCT
CTTTGCCCTCTGCTGGTTCCCTCTCAACTGCTATGTCCTCCTCTTGTCCAGCAAGGCCATCCACACCAACAATGC
CCTCTACTTTGCCTTCCACTGGTTTGCCATGAGCAGTACTTGTTATAACCCCTTCATCTACTGCTGGCTCAATGA
GAACTTTAGGGTTGAGCTTAAGGCATTGCTGAGCATGTGCCAAAGGCCACCCAAGCCGCAGGAAGACAGGCTACC
CTCCCCAGTTCCTTCCTTCAGGGTGGCATGGACAGAGAAGAGCCATGGTCGGAGGGCTCCACTACCTAATCACCA
CTTGCCCTCTTCCCAGATCCAGTCTGGGAAGACAGATCTGTCATCTGTGGAACCCGTTGTGGCCATGAGTTAGGG
AAAGCTGGAAGTTGGTGGGGGAGGGTTCTTTCCTCTCACAATTGACCAGACACTAACAGAGTTGGAAAGTAACAC
AGAAGCAGTGAGATGCTTGGGTTCCTAGGAACCTGTCCAGCCCCATCTGATTTGCAAACTTTCTAGAAGATGCCA
TGAGGTGGTGTGTGTAGATCTTTGAGCAAGAGCTCTGGAAACCACCTCAGCTTCAACAGAGGCTGGTCCAGTCAA
CCACCTCCAATTGTGTAGCATCTGCCACCTTGCCCTTCCTACTGCTGAGCAACCACAGGGGGACTTGAGCCATAC
TATTGGTGGGCCTGCCCCACATGCTCAGAAAAGAACAGGCACAAAGGCTTTCTGAAGTCATTGGAACAGGAATAA
TCACACAGCTTCAGTGACCTTGGCTCTATCCATGACCAGACAGGACCCATTTTGGCTTCTTAAAAACAAAGAGAA
ATTAGTATTGCCACTTTGAAAAGTTCAGAAAAGTAAAGAAATGAGTTCAGCCCTCAATTTGTAAAAAAAGGAAAA
AAGAAAAAAAAAAGAAAAAGAAAGAAAAAAGCCTGTTAATATGCTGTAAATTTATCTGTAGCTTTGCCTTCTGTG
TGTGTACATTTGTACTTTTAAAATCCTGAACTACACGTGTCCATGTAGATTGTAATAATTAGCAAGAAACTGGAA
TATATCAGAGTATTATTGAATTC (SEQ ID NO:13)

MKVPPVLLLFLLSSVRATEQPQVVTEHPSMEAALTGPNASSHFWANYTFSDWQNFVGRRRYGAESQNPTVKALLI
VAYSFTIVFSLFGNVLVCHVIFKNQRMHSATSLFIVNLAVADIMITLLNTPFTLVRFVNSTWVFGKGMCHVSRFA
QYCSLHVSALTLTAIAVDRHQVIMHPLKPRISITKGVIYIAVIWVMATFFSLPHAICQKLFTFKYSEDIVRSLCL
PDFPEPADLFWKYLDLATFILLYLLPLFIISVAYARVAKKLWLCNTIGDVTTEQYLALRRKKKTTVKMLVLVVVL
FALCWFPLNCYVLLLSSKAIHTNNALYFAFHWFAMSSTCYNPFIYCWLNENFRVELKALLSMCQRPPKPQEDRLP
SPVPSFRVAWTEKSHGRRAPLPNHHLPSSQIQSGKTDLSSVEPVVAMS (SEQ ID NO:14)

FIGURE 7

Underlined = deleted in targeting construct

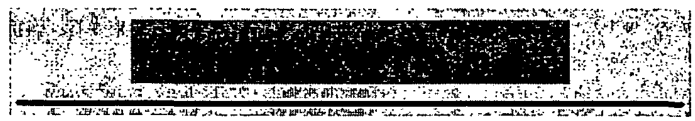
Bold = sequence flanking Neo insert in targeting construct

GGGGTGGCAGTCGGCACCATCAGGCTCCCTTGGCGTTTCGGAGTTTTCTCTGTGGTCCCG
ACTCTCCGGAGGATCTCGGTTGTCTCCCAAGTCGGAACCTGGCACGGTCCAGGTTCACTC
GGAGGTCCGGGCTTCCTCTGTGCCCCGTGCCCCTCGCTCCCAGGCTCCCTCTGTGGTGTG
GACTCCTCTAGCCCGGTGCGCTCAGCCCCTCGCACCCAGCCTCCAGGCACAGAGCCCGGC
AGGGAGCTCAGCCCTTGTGCCTAGAGCTGCAGTGGCTGGACATGAAGGTTCCTCCTGTCC
TGCTTCTCTTTCTTCTGTCCTCAGTGCGAGCTACTGAGCAACCGCAGGTCGTCACTGAGC
ATCCCAGCATGGAGGCAGCCCTGACCGGGCCCAACGCCTCCTCGCACTTCTGGGCCAACT
ACACTTTCTCTGACTGGCAGAACTTCGTGGGCAGGAGACGTTATGGGGCCGAGTCCCAGA
ACCCCACGGTGAAAGCACTGCTCATCGTGGCCTACTCATTCACCATCGTCTTCTCGCTCT
TCGGTAATGTCCTGGTCTGTCATGTCATCTTCAAGAACCAGCGCATGCACTCGGCCACCA
GCCTCTTCATTGTCAACCTGGCAGTGGCGGACATCATGATCACATTGCTCAACACGCCCT
TCACTTTGGTCCGCTTTGTGAACAGCACATGGGTGTTTGGGAAGGGCATGTGTCATGTCA
GTCGCTTTGCTCAGTACTGTTCTCTACATGTCT**CAGCACTGACTCTGACAGCTATCGCAG**
TGGACCGCCACCAGGTCATCATGCATCCACTGAAGCCTCGGATCTCCATCACCAAGGGTG
TCATATATATTGCTGTCATCTGGGTCATGGCTACCTTCTTCTCTCTGCCACATGCCATCT
GCCAGAAACTGTTTACCTTCAAGTACAGTGAGGACATTGTGCGCTCCCTCTGCCTGCCGG
ACTTCCCGGAGCCAGCTGACCTCTTCTGGAAGTATCTGGACCTGGCCACCTTCATCCTGC
TCTACCTACTTCCACTCTTCATTATCTCAGTGGCCTATGCTCGTGTGGCCAAGAAGCTGT
GGCTCTGTAACACCATTGGCGACGTGACCACAGAGCAGTACCTCGCCCTGCGACGCAAGA
AGAAGACCACCGTGAAGATGCTGGTGCTTGTGGTAGTCCTCTTTGCCCTCTGCTGGTTCC
CTCTCAACTGCTATGTCCTCCTCTTGTCCAGCAAGGCCATCCACACCAACAATGCCCTCT
ACTTTGCCTTCCACTGGTTTGCCATGAGCAGTACTTGTTATAACCCCTTCATCTACTGCT
GGCTCAATGAGAACTTTAGGGTTGAGCTTAAGGCATTGCTGAGCATGTGCCAAAGGCCAC
CCAAGCCGCAGGAAGACAGGCTACCCTCCCCAGTTCCTTCCTTCAGGGTGGCATGGACAG
AGAAGAGCCATGGTCGGAGGGCTCCACTACCTAATCACCACTTGCCCTCTTCCCAGATCC
AGTCTGGGAAGACAGATCTGTCATCTGTGGAACCCGTTGTGGCCATGAGTTAGGGAAAGC
TGGAAGTTGGTGGGGGAGGGTTCTTTCCTCTCACAATTGACCAGACACTAACAGAGTTGG
AAAGTAACACAGAAGCAGTGAGATGCTTGGGTTCCTAGGAACCTGTCCAGCCCCATCTGA
TTTGCAAACTTTCTAGAAGATGCCATGAGGTGGTGTGTGTAGATCTTTGAGCAAGAGCTC
TGGAAACCACCTCAGCTTCAACAGAGGCTGGTCCAGTCAACCACCTCCAATTGTGTAGCA
TCTGCCACCTTGCCCTTCCTACTGCTGAGCAACCACAGGGGGACTTGAGCCATACTATTG
GTGGGCCTGCCCCACATGCTCAGAAAAGAACAGGCACAAAGGCTTTCTGAAGTCATTGGA
ACAGGAATAATCACACAGCTTCAGTGACCTTGGCTCTATCCATGACCAGACAGGACCCAT
TTTGGCTTCTTAAAAACAAAGAGAAATTAGTATTGCCACTTTGAAAAGTTCAGAAAAGTA
AAGAAATGAGTTCAGCCCTCAATTTGTAAAAAAAGGAAAAAAGAAAAAAAAAAGAAAAAG
AAAGAAAAAAGCCTGTTAATATGCTGTAAATTTATCTGTAGCTTTGCCTTCTGTGTGTGT
ACATTTGTACTTTTAAAATCCTGAACTACACGTGTCCATGTAGATTGTAATAATTAGCAA
GAAACTGGAATATATCAGAGTATTATTGAATTC

FIGURE 8A

Gene Sequence Structure * 299 bp Sequence Deleted 753 bp

Size of partial
cDNA: 2253 bp

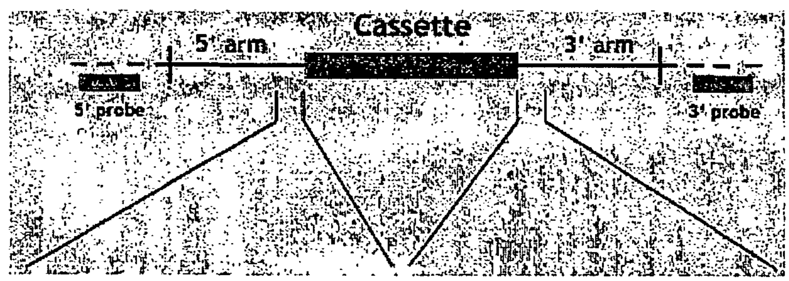


Targeting Vector*
(genomic sequence)

Construct Number: 463

Arm Length:
5': 2.5 kb
3': 0.6 kb

LacZ-Neo



* Not drawn to scale

5'>CTGGCACGGTCCAGGTTCACT CGGAGGCCCGGGCTTCCTCTGTGC CCCGTGCCCCTCGCTCCCTGGCTC CCTCTGTGGTGTGGACTCCTCTAG CCCGGTGCGCTCAGCCCCTCGCAC CCAGCCTCCAGGCACAGAGCCCGG CAGGGAGCTCAGCCCTTGTGCCTA GAGCTGCAGTGGCTGGACATGAAG GTTTCTCCTGT<3' (SEQ ID NO:15)	5'>CAGCACTGACTCTGACAGCTA TCGCAGTGGACCGCCACCAGGTGA GAGCACCTGTCCCCAGCAGCATGC TCCCATCTCCGTCTATGCCTGGCT GGCTGGTGGGAATACTGCCACCAC GGTCTGTAGGGAATACTCTCAGGA CAGTGACTCATTCAGTCCCGCTGA CAGCGTGTGTGCTTGCCTCCTTGT TGATCAATTTG<3' (SEQ ID NO:16)

FIGURE 8B

AAGCTTGTTCACAGGTGGCCATGACCTCACTGCCCCCTGGAACCACTGGGGACCCCGATTTGTTTTCTGGGCCGT
CGCCAGCCGGCTCCACTCCAGCCAACCAGAGTGCAGAGGCTTCAGAGAGCAATGTGTCTGCGACGGTTCCCAGAG
CTGCAGCAGTCACGCCGTTCCAGAGCCTGCAACTAGTGCACCAGCTGAAGGGACTGATCGTGATGCTGTACAGCA
TCGTGGTGGTCGTGGGTCTGGTGGGCAACTGCCTTCTTGTGCTGGTGATCGCGCGCGTGCGCCGGCTGCACAACG
TGACCAACTTCCTCATCGGCAACCTGGCCTTGTCCGATGTGCTCATGTGTGCCGCCTGTGTGCCTCTCACGCTGG
CCTACGCCTTTGAACCTCGTGGCTGGGTGTTCGGTGGAGGCCTGTGCCACCTTGTTTTCTTCCTGCAGCCGGTCA
CCGTCTACGTATCGGTGTTCACACTCACCACAATCGCTGTGGACCGCTATGTGGTTCTGGTGCACCCGCTACGTC
GGCGCATTTCACTGAAGCTCAGCGCCTACGCTGTGCTGGGCATCTGGGCTCTATCTGCAGTGCTGGCGCTGCCGG
CCGCGGTGCACACCTACCATGTAGAGCTCAAGCCCCACGACGTGCGCCTCTGCGAGGAGTTCTGGGGTTCGCAGG
AGCGCCAGAGACAGATCTATGCCTGGGGGCTGCTGCTGGGCACCTATTTGCTCCCCCTGCTGGCCATTCTCCTGT
CTTACGTCCGGGTGTCGGTGAAGTTGCGGAACCGCGTGGTGCCTGGCAGCGTGACCCAGAGCCAGGCTGACTGGG
ACCGAGCGCGTCGCCGTCGCACTTTCTGCCTGCTGGTGGTGGTGGTGGTCGTGTTCGCGCTCTGCTGGCTGCCTC
TGCACATTTTCAACCTTCTGCGGGACCTGGACCCGCGTGCCATCGACCCCTACGCCTTCGGGCTGGTGCAGCTCC
TCTGCCACTGGCTTGCCATGAGCTCCGCCTGCTACAACCCCTTCATCTATGCGTGGCTGCACGACAGCTTCCGAG
AGGAGCTACGCAAGATGCTTCTGTCTTGGCCCCGCAAGATCGTGCCTCATGGCCAGAATATGACCGTCAGTGTGG
TCATCTGATGACACTCTGCCAGGTCTGAGCTGGGGAGCTTCAGAGTCATTAGCCCCGGAAGTCATCATGAGATCT
TATTCCCATTTGCCAACACAAGTGAAATTCCTACCCAGCCCTCTAGACTAGCTATATGTCTGTTTGTGTGCTTGT
GTGAAATGTTAGATAGGCTTTGGAGTAGAAGAGCCTGAGGAGAGAGAAGCAGATTTATCAGGTCTACCCCTGAAA
GCAAAAGAGATATTTTTCTGGGTTCAGAAAAAATAAATCAGAAGCAGTCTTTCTTTCCTTCTCTCTCTCTCTC
TCTCTCTCTCTCTCTNNCTTTCTTTCT (SEQ ID NO:17)

MTSLPPGTTGDPDLFSGPSPAGSTPANQSAEASESNVSATVPRAAAVTPFQSLQLVHQLKGLIVMLYSIVVVVGL
VGNCLLVLVIARVRRLHNVTNFLIGNLALSDVLMCAACVPLTLAYAFEPRGWVFGGGLCHLVFFLQPVTVYVSVF
TLTTIAVDRYVVLVHPLRRRISLKLSAYAVLGIWALSAVLALPAAVHTYHVELKPHDVRLCEEFWGSQERQRQIY
AWGLLLGTYLLPLLAILLSYVRVSVKLRNRVVPGSVTQSQADWDRARRRRTFCLLVVVVVVFALCWLPLHIFNLL
RDLDPRAIDPYAFGLVQLLCHWLAMSSACYNPFIYAWLHDSFREELRKMLLSWPRKIVPHGQNMTVSVVI (SEQ ID NO:18)

FIGURE 9

<u>Underlined</u> = deleted in targeting construct

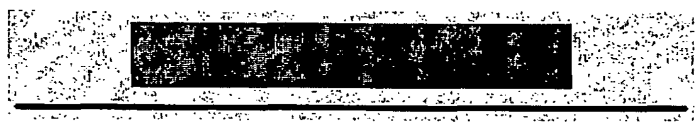
Bold = sequence flanking Neo insert in targeting construct

AAGCTTGTTCACAG**GTGGCCATGACCTCACTGCCCCCTGGAACCACTGGGGACCCCGATT**
TGTTTTCTGGGCCGTCGCCAGCCGGCTCCACTCCAGCCAACCAGAGTGCAGAGGCTTCAG
AGAGCAATGTGTCTGCGACGGTTCCCAGAGCTGCAGCAGTCACGCCGTTCCAGAGCCTGC
AACTAGTGCACCAGCTGAAGGGACTGATCGTGATGCTGTACAGCATCGTGGTGGTCGTGG
GTCTGGTGGGCAACTGCCTTCTTGTGCTGGTGA<u>TCGCGCGCGTGCGCCGGCTGCACAACG</u>
<u>TGACCAACTTCCTCATCGGCAACCTGGCCTTGTCCGATGTGCTCATGTGTGCCGCCTGTG</u>
<u>TGCCTCTCACGCTGGCCTACGCCTTTGAACCTCGTGGCTGGGTGTTCGGTGGAGGCCTGT</u>
<u>GCCACCTTGTTTTCTTCCTGCAGCCGGTCACCGTCTACGTAT</u>**CGGTGTTCACACTCACCA**
CAATCGCTGTGGACCGCTATGTGGTTCTGGTGCACCCGCTACGTCGGCGCATTTCACTGA
AGCTCAGCGCCTACGCTGTGCTGGGCATCTGGGCTCTATCTGCAGTGCTGGCGCTGCCGG
CCGCGGTGCACACCTACCATGTAGAGCTCAAGCCCCACGACGTGCGCCTCTGCGAGGAGT
TCTGGGGTTCGCAGGAGCGCCAGAGACAGATCTATGCCTGGGGGCTGCTGCTGGGCACCT
ATTTGCTCCCCCTGCTGGCCATTCTCCTGTCTTACGTCCGGGTGTCGGTGAAGTTGCGGA
ACCGCGTGGTGCCTGGCAGCGTGACCCAGAGCCAGGCTGACTGGGACCGAGCGCGTCGCC
GTCGCACTTTCTGCCTGCTGGTGGTGGTGGTGGTCGTGTTCGCGCTCTGCTGGCTGCCTC
TGCACATTTTCAACCTTCTGCGGGACCTGGACCCGCGTGCCATCGACCCCTACGCCTTCG
GGCTGGTGCAGCTCCTCTGCCACTGGCTTGCCATGAGCTCCGCCTGCTACAACCCCTTCA
TCTATGCGTGGCTGCACGACAGCTTCCGAGAGGAGCTACGCAAGATGCTTCTGTCTTGGC
CCCGCAAGATCGTGCCTCATGGCCAGAATATGACCGTCAGTGTGGTCATCTGATGACACT
CTGCCAGGTCTGAGCTGGGGAGCTTCAGAGTCATTAGCCCCGGAAGTCATCATGAGATCT
TATTCCCATTTGCCAACACAAGTGAAATTCCTACCCAGCCCTCTAGACTAGCTATATGTC
TGTTTGTGTGCTTGTGTGAAATGTTAGATAGGCTTTGGAGTAGAAGAGCCTGAGGAGAGA
GAAGCAGATTTATCAGGTCTACCCCTGAAAGCAAAAGAGATATTTTTCTGGGTTCAGAAA
AAATAAATCAGAAGCAGTCTTTCTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
NNCTTTCTTTCT

FIGURE 10A

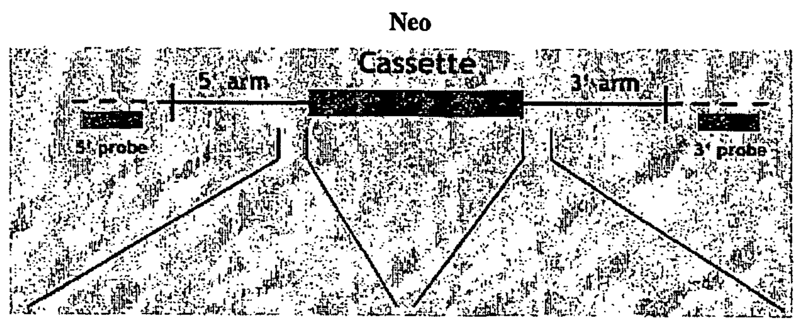
Gene Sequence Structure * 274 bp Sequence Deleted 462 bp

Size of full-length cDNA: 1452 bp



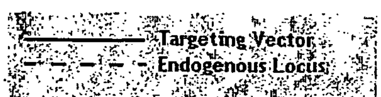
Targeting Vector*
(genomic sequence)

Construct Number: 262



Arm Length:
5': 5.3 kb
3': 1 kb

5'>GTTGCCAGCCAGCTCCACTCC AGCCAACCAGAGCGCAGAGGCCTC GGAGGGCAACCTGTCTGCGACGGT TCCCCGAGCTGCAGCAGTCACGCC GTTCCAGAGCCTGCAGCTGGTGCA CCAGCTGAAGGGGTTGATCGTGAT GCTGTACAGCATTGTGGTGGTCGT GGGTCTGGTGGGCAACTGCCTTCT TGTGCTGGTGA<3' (SEQ ID NO:19)	5'>CGGTGTTCACACTCACCACAA TCGCCTTGGACCGCTATGTGGTTC TGGTGCACCCGCTACGTCGGTGCA TTTCACTGAGGCTCAGCGCCTACG CGGTGCTGGGCATCTGGGCTCTAT CTGCAGTGCTGGCGCTGCCGGCCG CGGTGCACACCTACCATGTGGAGC TCAAGCCCCACGACGTGCGCCTCT GCGAGGAGTTC<3' (SEQ ID NO:20)



* Not drawn to scale

FIGURE 10B

GGGACAGTGGGTCCCAATGGCTCTAGGGTCCTCCTGTGTAGCTGGGGAGATAACAAAAAAGGGATTCTTTTGAGG
CTTCCAACAGGATATAGGACCTGGTGAGCCTTTGTCTCTCTGCATAGGGACAGTGACTGTGTCCATCACAGAGGC
TGTTTAGGGCATAGAAGTAGGTTACTGCCTTGAACCTCTGACACTAATCTTTTCCCACAGGACAAGTTTCCCACG
GGCTCTCCTCACTGAGCAGTGGTTCTCCCCCTGGAATCCCAGTGTGAGGACCGAGATGGCTCTGAGCCTGGAGTC
TACAACAAGCTTTCATATGCTCACCGTGTCCGGAAGCACTGTGACTGAGCTGCCTGGTGACTCCAACGTGTCCCT
CAACAGTTCCTGGTCCGGCCCAACAGATCCCAGCTCCCTGAAAGACCTTGTGGCCACGGGTGTCATCGGGGCAGT
GCTCTCAGCCATGGGTGTGGTGGGCATGGTGGGAAATGTATACACTTTGGTGGTCATGTGCCGGTTTCTGCGTGC
CTCGGCCTCCATGTACGTCTATGTGGTCAACCTAGCGCTGGCTGATCTGCTGTACCTGCTGAGCATTCCCTTCAT
CATAGCCACCTACGTCACTAAGGACTGGCACTTTGGAGATGTGGGCTGCAGAGTCCTCTTTAGCCTGGACTTCCT
GACAATGCACGCCAGCATCTTCACCCTGACCATAATGAGCAGCGAACGCTATGCAGCCGTACTGAGGCCTCTGGA
CACAGTCCAGCGCTCCAAGGGTTACCGTAAGCTGCTGGTGCTGGGCACCTGGTTGCTGGCACTGCTGCTGACCCT
ACCCATGATGCTTGCCATCCAGCTGGTCCGCAGGGGCTCTAAGAGCCTCTGCCTGCCAGCCTGGGGCCCTCGTGC
CCACCGTACTTACCTAACGTTGCTCTTTGGGACCAGCATTGTGGGGCCTGGCTTGGTCATTGGGCTGCTCTATGT
CCGTCTGGCCAGGGCCTACTGGCTATCTCAGCAAGCTTCTTTCAAGCAGACACGGCGGCTGCCCAACCCCAGGGT
GCTCTACCTCATCCTTGGTATCGTCCTTCTCTTCTGGGCCTGCTTTCTACCCTTCTGGCTGTGGCAGCTGCTGGC
CCAGTACCACGAGGCCATGCCACTGACTCCCGAGACTGCACGCATTGTCAACTACCTGACCACCTGCCTCACTTA
TGGCAACAGTTGCATCAATCCCTTGCTCTACACTCTGCTCACCAAGAACTATCGAGAGTACCTACGTGGCCGCCA
GCGGTCACTGGGTAGTAGTTGCCACAGCCCAGGGAGTCCTGGCAGCTTCCTGCCCAGCCGAGTCCACCTCCAGCA
GGACTCGGGCCGCTCGCTGTCCTCCAGCAGCCAACAGGCCACAGAGACCCTCATGCTGTCTCCAGTCCCCCGTAA
CGGGGCCCTTCTCTGAGAGTGCACTGTGCAATCCTGGCATAGGAAAGGACCCAAAGGCGTGCGGCTCCGGAGCGC
ATTTCCCAGAATCCCCTGCTCAAACCTAACTGGCTCGTC (SEQ ID NO:21)

MALSLESTTSFHMLTVSGSTVTELPGDSNVSLNSSWSGPTDPSSLKDLVATGVIGAVLSAMGVVGMVGNVYTLVV
MCRFLRASASMYVYVVNLALADLLYLLSIPFIIATYVTKDWHFGDVGCRVLFSLDFLTMHASIFTLTIMSSERYA
AVLRPLDTVQRSKGYRKLLVLGTWLLALLLTLPMMLAIQLVRRGSKSLCLPAWGPRAHRTYLTLLFGTSIVGPGL
VIGLLYVRLARAYWLSQQASFKQTRRLPNPRVLYLILGIVLLFWACFLPFWLWQLLAQYHEAMPLTPETARIVNY
LTTCLTYGNSCINPLLYTLLTKNYREYLRGRQRSLGSSCHSPGSPGSFLPSRVHLQQDSGRSLSSSSQQATETLM
LSPVPRNGALL (SEQ ID NO:22)

FIGURE 11

Underlined = deleted in targeting construct

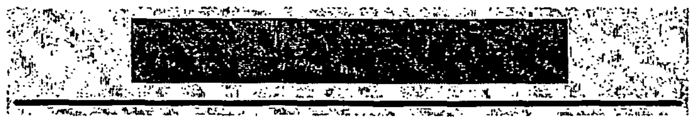
Bold = sequence flanking Neo insert in targeting construct

GGGACAGTGGGTCCCAATGGCTCTAGGGTCCTCCTGTGTAGCTGGGGAGATAACAAAAAA
GGGATTCTTTTGAGGCTTCCAACAGGATATAGGACCTGGTGAGCCTTTGTCTCTCTGCAT
AG**GGACAGTGACTGTGTCCATCACAGAGGCTGTTTAGGGCATAGAAGTAGGTTACTGCCT**
TGAACCTCTGACACTAATCTTTTCCCACAGGACAAGTTTCCCACGGGCTCTCCTCACTGA
GCAGTGGTTCTCCCCCTGGAATCCCAGTGTGAGGACCGAGATGGCTCTGAGCCTGGAGTC
TACAACAAGCTTTCATATGCTCACCGTGTCCGGAAGCACTGTGACTGAGCTGCCTGGTGA
CTCCAACGTGTCCCTCAACAGTTCCTGGTCCGGCCCAACAGATCCCAGCTCCCTGAAAGA
CCTTGTGGCCACGGGTGTCATCGGGGCAGTGCTCTCAGCCATGGGTGTGGTGGGCATGGT
GGGAAATGTATACACTTTGGTGGTCATGTGCCGGTTTCTGCGTGCCTCGGCCTCCATGTA
CGTCTATGTGGTCAACCTAGCGCTGGCTGATCTGCTGTACCTGCTGAGCATTCCCTTCAT
CATAGCCACCTACGTCACTAAGGACTGGCACTTTGGAGATGTGGGCTGCAGAGTCCTCTT
TAGCCTGGACTTCCTGACAATGCACGCCAGCATCTTCACCCTGACCATAATGAGCAGCGA
ACGCTATGCAGCCGTACTGAGGCCTCTGGACACAGTCCAGCGCTCCAAGGGTTACCGTAA
GCTGCTGGTGCTGGGCAC**CTGGTTGCTGGCACTGCTGCTGACCCTACCCATGATGCTTGC**
CATCCAGCTGGTCCGCAGGGGCTCTAAGAGCCTCTGCCTGCCAGCCTGGGGCCCTCGTGC
CCACCGTACTTACCTAACGTTGCTCTTTGGACCAGCATTGTGGGGCCTGGCTTGGTCAT
TGGGCTGCTCTATGTCCGTCTGGCCAGGGCCTACTGGCTATCTCAGCAAGCTTCTTTCAA
GCAGACACGGCGGCTGCCCAACCCCAGGGTGCTCTACCTCATCCTTGGTATCGTCCTTCT
CTTCTGGGCCTGCTTTCTACCCTTCTGGCTGTGGCAGCTGCTGGCCCAGTACCACGAGGC
CATGCCACTGACTCCCGAGACTGCACGCATTGTCAACTACCTGACCACCTGCCTCACTTA
TGGCAACAGTTGCATCAATCCCTTGCTCTACACT CTGCTCACCAAGAACTATCGAGAGTA
CCTACGTGGCCGCCAGCGGTCACTGGGTAGTAGTTGCCACAGCCCAGGGAGTCCTGGCAG
CTTCCTGCCCAGCCGAGTCCACCTCCAGCAGGACTCGGGCCGCTCGCTGTCCTCCAGCAG
CCAACAGGCCACAGAGACCCTCATGCTGTCTCCAGTCCCCCGTAACGGGGCCCTTCTCTG
AGAGTGCACTGTGCAATCCTGGCATAGGAAAGGACCCAAAGGCGTGCGGCTCCGGAGCGC
ATTTCCCAGAATCCCCTGCTCAAACCTAACTGGCTCGTC

FIGURE 12A

Gene Sequence Structure * 566 bp Sequence Deleted 798 bp

Size of CDS: 1539 bp

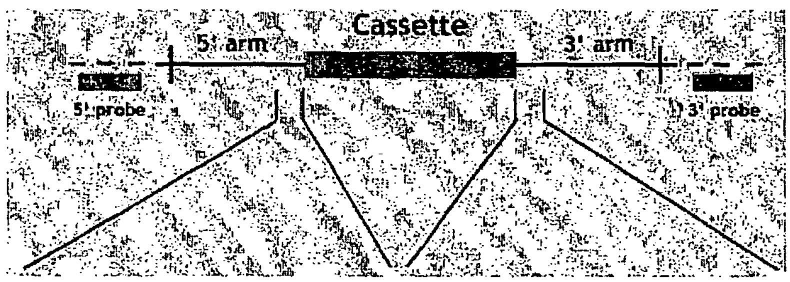


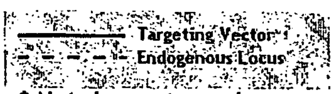
Targeting Vector*
(genomic sequence)

Construct Number: 283

Arm Length:
5': 0.6 kb
3': 5 kb

Neo





* Not drawn to scale

5'>ATGTGTCCCACAACAGTTCCT
GGACTGGCCCAACAGATCCCAGCT
CCCTGCAAGACCTTGTGGCCACCG
GTGTCATCGGGGCAGTGCTCTCAA
CCATGGGTGTGGTGGGCGTGGTGG
GCAATGTGTACACTCTGGTGGTCA
TGTGCCGATTTCTGCGTGCCTCGG
CCTCCATGTACGTCTATGTGGTCA
ACCTAGCGCTG<3'
(SEQ ID NO:23)

5'>CTGGTTGCTGGCACTGCTGCT
GACCTTACCCATGATGCTTGCCAT
CCGGCTGGTCCGTAGGGGCTCTAA
GAGCCTCTGCCTGCCAGCCTGGGG
CCCTCGTGCCCACCGTACCTACCT
GACGCTGCTCTTTGGGACCAGCAT
TGTGGGGCCTGGCCTGGTCATTGG
GCTGCTCTATATCCGTCTGGCCAG
GGCCTATTGGC<3'
(SEQ ID NO:24)

FIGURE 12B

AAGCTTCCTTGAGGTTTCTAAAATTTATACAAAAACATCATATGTAAGTAAACTCACCAGATTTGGCATCTGCTC
TTTGGTGATGGACCCAGAAGAAACTTCAGTTTATTTGGATTATTACTATGCTACGAGCCCAAACTCTGACATCAG
GGAGACCCACTCCCATGTTCCTTACACCTCTGTCTTCCTTCCAGTCTTTTACACAGCTGTGTTCCTGACTGGAGT
GCTGGGGAACCTTGTTCTCATGGGAGCGTTGCATTTCAAACCCGGCAGCCGAAGACTGATCGACATCTTTATCAT
CAATCTGGCTGCCTCTGACTTCATTTTTCTTGTCACATTGCCTCTCTGGGTGGATAAAGAAGCATCTCTAGGACT
GTGGAGGACGGGCTCCTTCCTGTGCAAAGGGAGCTCCTACATGATCTCCGTCAATATGCACTGCAGTGTCCTCCT
GCTCACTTGCATGAGTGTTGACCGCTACCTGGCCATTGTGTGGCCAGTCGTATCCAGGAAATTCAGAAGGACAGA
CTGTGCATATGTAGTCTGTGCCAGCATCTGGTTTATCTCCTGCCTGCTGGGGTTGCCTACTCTTCTGTCCAGGGA
GCTCACGCTGATTGATGATAAGCCATACTGTGCAGAGAAAAAGGCAACTCCAATTAAACTCATATGGTCCCTGGT
GGCCTTAATTTTCACCTTTTTTGTCCCTTTGTTGAGCATTGTGACCTGCTACTGTTGCATTGCAAGGAAGCTGTG
TGCCCATTACCAGCAATCAGGAAAGCACAACAAAAAGCTGAAGAAATCTATAAAGATCATCTTTATTGTCGTGGC
AGCCTTTCTTGTCTCCTGGCTGCCCTTCAATACTTTCAAGTTCCTGGCCATTGTCTCTGGGTTGCGGCAAGAACA
CTATTTACCCTCAGCTATTCTTCAGCTTGGTATGGAGGTGAGTGGACCCTTGGCATTTGCCAACAGCTGTGTCAA
CCCTTTCATTTACTATATCTTCGACAGCTACATCCGCCGGGCCATTGTCCACTGCTTGTGCCCTTGCCTGAAAAA
CTATGACTTTGGGAGTAGCACTGAGACATCAGATAGTCACCTCACTAAGGCTCTCTCCACCTTCATTCATGCAGA
AGATTTTGCCAGGAGGAGGAAGAGGTCTGTGTCACTCTAAAGGGAACTGTGACATTTCAAGCTCTGTTGGTGGGT
TTAGGAGTTAATTTTTGTCAGCAACAAAGAAA (SEQ ID NO:25)

MDPEETSVYLDYYYATSPNSDIRETHSHVPYTSVFLPVFYTAVFLTGVLGNLVLMGALHFKPGSRRLIDIFIINL
AASDFIFLVTLPLWVDKEASLGLWRTGSFLCKGSSYMISVNMHCSVLLLTCMSVDRYLAIVWPVVSRKFRRTDCA
YVVCASIWFISCLLGLPTLLSRELTLIDDKPYCAEKKATPIKLIWSLVALIFTFFVPLLSIVTCYCCIARKLCAH
YQQSGKHNKKLKKSIKIIFIVVAAFLVSWLPFNTFKFLAIVSGLRQEHYLPSAILQLGMEVSGPLAFANSCVNPF
IYYIFDSYIRRAIVHCLCPCLKNYDFGSSTETSDSHLTKALSTFIHAEDFARRRKRSVSL (SEQ ID NO:26)

FIGURE 13

Underlined = deleted in targeting construct

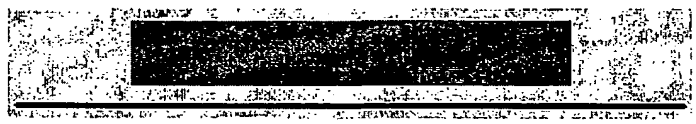
Bold = sequence flanking Neo insert in targeting construct

AAGCTTCCTTGAGGTTTCTAAAATTTATACAAAAACATCATATGTAAGTAAACTCACCAG
ATTTGGCATCTGCTCTTTGGTGATGGACCCAGAAGAAACTTCAGTTTATTTGGATTATTA
CTATGCTACGAGCCCAAACTCTGACATCAGGGAGACCCACTCCCATGTTCCTTACACCTC
TGTCTTCCTTCCAGTCTTTTACACAGCTGTGTTCCTGACTGGAGTGCTGGGGAACCTTGT
TCTCATGGGAGCGTTGCATTTCAAACCCGGCAGCCGAAGACTGATCGACATCTTTATCAT
CAATCTGGCTGCCTCTGACTTCATTTTTCTTGTCACATTGCCTCTCTGGGTGGATAAAGA
AGCATCTCTAGGACTGTGGAGGACGGGCTCCTTCCTGTGCAAAGGGAGCTCCTACATGAT
CTCCGTCAATATGCACTGCAGTGTCCTCCTGCTCACTTGCATGAGTGTTGACCGCTACCT
GGCCATTGTGTGGCCAGTCGTATCCAGGAAATTCAGAAGGACAGACTGTGCATATGTAGT
CTGTGCCAGCATCTGGTTTATCTCCTGCCTGCTGGGGTTGCCTACTCTTCTGTCCAGGGA
GCTCACGCTGATTGATGATAAGCCATACTGTGCAGAGAAAAAGGCAACTCCAATTAAACT
CATATGGTCCCTGGTGGCCTTAATTTTCACCTTTTTTGTCCCTTTGTTGAGCATTGTGAC
CTGCTACTGTTGCATTGCAAGGAAGCTGTGTGCCCATTACCAGCAATCAGGAAAGCACAA
CAAAAAGCTGAAGAAATCTATAAAGATCATCTTTATTGTCGTGGCAGCCTTTCTTGTCTC
CTGGCTGCCCTTCAATACTTTCAAGTTCCTGGCCATTGTCTCTGGGTTGCGGCAAGAACA
CTATTTACCCTCAGCTATTCTTCAGCTTGGTATGGAGGTGAGTGGACCCTTGGCATTTGC
CAACAGCTGTGTCAACCCTTTCATTTACTATATCTTCGACAGCTACATCCGCCGGGCCAT
TGTCCACTGCTTGTGCCCTTGCCTGAAAAACTATGACTTTGGGAGTAGCACTGAGACATC
AGATAGTCACCTCACTAAGGCTCTCTCCACCTTCATTCATGCAGAAGATTTTGCCAGGAG
GAGGAAGAGGTCTGTGTCACTCTAAAGGGAACTGTGACATTTCAAGCTCTGTTGGTGGGT
TTAGGAGTTAATTTTTGTCAGCAACAAAGAAA

FIGURE 14A

Gene Sequence Structure * 448 bp Sequence Deleted 746 bp

Size of CDS: 1232 bp

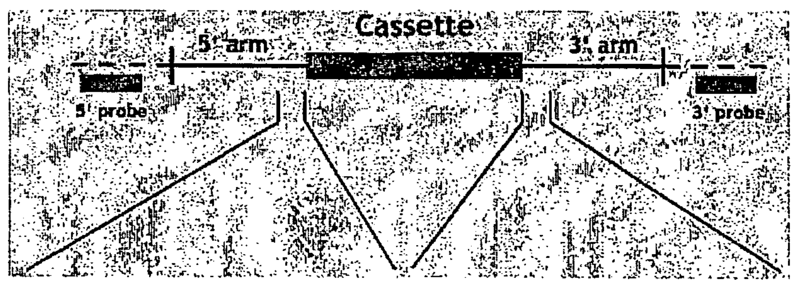


Targeting Vector*
(genomic sequence)

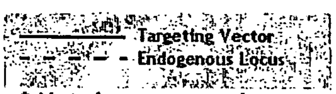
Construct Number: 2454

Arm Length:
5': 3.2 kb
3': 1 kb

LacZ-Neo



5'>ATAGCTCTGCATTTCAAACGCGGCAACCGAAGATTGATCGACATCTTTATCATCAACCTGGCTGCCTCTGACTTCATTTTCCTTGTCACAGTGCCTCTTTGGATGGATAAGGAAGCCTCTCTAGGATTATGGAGGACTGGCTCTTTCCTGTGCAAAGGCAGCTCCTATGTGATCTCCGTGAACATGCACTGTAGTGTCTT<3' (SEQ ID NO:27)	5'>TGTGTGCTCATTACCAGCAGTCGGGAAAGCATAACAAGAAACTAAAGAAGTCCATAAAGATCGTTATTATTGCGGTGGCGGCCTTCACCGTCTCCTGGGTACCCTTTAACACTTTCAAGCTCCTAGCCATTGTTTCAGGGTTCCAGCCAGAAGGCCTTTTTCACTCCGAGGCTCTGCAGCTGGCCATGAATGTGACTGGG<3' (SEQ ID NO:28)



* Not drawn to scale

FIGURE 14B

GAGAAGTCAAACTTGGTATACCGGGAGAGGGAGAGATGAAAACAACTTCTCAGTTTCTCTGTGGCTGACAGAAGC
TCCTGAGAGAGGCTTTCAGAACTTGATTCCAGCTGGCCCGTGAAAGCTGTGTCTGGACGGGAGAGCCTCAGAGGA
ATAAACAGTGCCTTCTCTGGCTTCTCCGGGTCAAGACCAGTCTGTTCAGAAGGAAGTGGTGTTTCTCTGACCCCA
CTGTTCCCACAGCCCCCTGGATGGTAGATCATGGGGCTGCACCCTTCTCCACCATTGGGAACTCAGTCAAACGGT
TATCTTTATCTCGCCAGCTTGTCCTCGTGCTTGACCAGCGCAGTTTCACTTTTGCAAACATCTGTTTATATCCCT
TGAGAGAAAAATATCAAGCAACCTGCCTCAAACGACCCTGTTTTGGCCGGTGAGCAAGGACAGCCTCCGATGGAT
AACTACACAGTGGCCCCGGACGATGAATATGATGTCCTAATCTTAGACGACTACCTGGACAACAGTGGGCCGGAC
CAAGTTCCGGCCCCCGAGTTCCTCTCCCCCCAGCAGGTGCTGCAGTTCTGCTGCGCGGTGTTTGCGGTGGGTCTC
TTGGACAACGTGCTGGCGGTGTTTATCTTGGTGAAATACAAAGGACTCAAGAATCTGGGGAACATCTACTTCCTA
AACCTGGCACTTTCAAACCTGTGTTTCCTGCTTCCCCTGCCGTTCTGGGCCCATACTGCAGCACACGGGGAAAGC
CCTGGCAACGGGACCTGTAAAGTTCTTGTCGGACTCCACTCCTCGGGCTTATACAGCGAGGTGTTTTCCAACATC
CTCCTCCTTGTGCAAGGATACAGGGTGTTTTCCCAAGGGCGACTGGCCTCCATCTTCACGACAGTGTCTTGTGGT
ATTGTTGCGTGCGTCCTGGCATGGGCCATGGCTACTGCGCTCTCTTTGCCCGAGTCTGTGTTTTATGAGCCTCGG
ATGGAAAGACAGAAACACAAGTGTGCCTTTGGCAAACCTCACTTCTTGCCAATCGAAGCGCCGCTCTGGAAGTAC
GTTCTGACGTCAAAAATGATCATCTTGGTACTTGCTTTTCCTCTGCTGGTTTTTATAATCTGCTGCAGGCAACTG
AGGAGAAGGCAGAGCTTCAGGGAGAGACAGTACGACCTCCACAAGCCGGCTCTTGTCATAACGGGCGTGTTCCTT
TTGATGTGGGCGCCTTACAACACTGTGCTTTTCCTGTCTGCTTTCCAGGAACACTTGTCCCTGCAGGATGAGAAG
AGCAGCTACCACCTGGACGCAAGTGTTCAGGTCACACAGCTGGTAGCGACCACCCACTGCTGCGTCAACCCGCTG
CTCTATTTGCTTCTTGACCGGAAGGCCTTTATGAGATACCTTCGCAGCCTGTTCCCACGGTGCAATGATATCCCC
TATCAAAGTAGTGGAGGCTCTCAGCAAGCGCCGCCAAGGAAAGGTCATGGCAGGCCCATTGAACTGTACAGCAAT
TTGCATCAAAGGCAGGATATAATATAATATAAACACTCAACCTTGTCTCTTCACATTGTTTTATGCCATTTTTAT
ATGTTTGTATAAAAATAGAATATAAGAGGAAAGGGGGCTGGTTAATGAACATTTACTACTAAGCGTTGAATTTGT
CTCTGGCACACGGTTAAGTTATTTATAAAGGTGAATTCCACTCTTCCTCTTTAGCGTTTGTCCGCAGTGTAGGCG
TGGGGCGGTGTGCATTCCATAAACTGAAGCTTGGTAGTTGTCCAAAGTTACATAATTATAAAGTGGCAAAACTGA
TGTGTAAACTCAGAGATCACTGGCCCCAGAGTCTGTGAGTTATCCAAGCGTGCCAGCCTTTCAGTAACTGTCAGC
TAGCAGTGGTGGCATGTCTGTCATCCCAGCACTCTAGGAGGCGGAGGCAGGAGGCATCTCACAAATGCAAGTATA
ACCTGACTTCCCCCAACCTCAATTCTTTTTTTTTTGTTGTTGTTGTTCATCAAAATACATGCAATAAATATATA
TATTTTTAAAGCAAAAAAAA (SEQ ID NO:29)

MDNYTVAPDDEYDVLILDDYLDNSGPDQVPAPEFLSPQQVLQFCCAVFAVGLLDNVLAVFILVKYKGLKNLGNIY
FLNLALSNLCFLLPLPFWAHTAAHGESPGNGTCKVLVGLHSSGLYSEVFSNILLLVQGYRVFSQGRLASIFTTVS
CGIVACVLAWAMATALSLPESVFYEPRMERQKHKCAFGKPHFLPIEAPLWKYVLTSKMIILVLAFPLLVFIICCR
QLRRRQSFRERQYDLHKPALVITGVFLLMWAPYNTVLFLSAFQEHLSLQDEKSSYHLDASVQVTQLVATTHCCVN
PLLYLLLDRKAFMRYLRSLFPRCNDIPYQSSGGSQQAPPRKGHGRPIELYSNLHQRQDII (SEQ ID NO:30)

FIGURE 15

<u>Underlined</u> = deleted in targeting construct

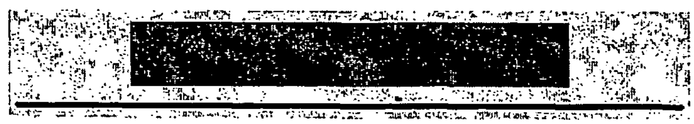
Bold = sequence flanking Neo insert in targeting construct

GAGAAGTCAAACTTGGTATACCGGGAGAGGGAGAGATGAAAACAACTTCTCAGTTTCTCT
GTGGCTGACAGAAGCTCCTGAGAGAGGCTTTCAGAACTTGATTCCAGCTGGCCCGTGAAA
GCTGTGTCTGGACGGGAGAGCCTCAGAGGAATAAACAGTGCCTTCTCTGGCTTCTCCGGG
TCAAGACCAGTCTGTTCAGAAGGAAGTGGTGTTTCTCTGACCCCACTGTTCCCACAGCCC
CCTGGATGGTAGATCATGGGGCTGCACCCTTCTCCACCATTGGGAACTCAGTCAAACGGT
TATCTTTATCTCGCCAGCTTGTCCTCGTGCTTGACCAGCGCAGTTTCACTTTTGCAAACA
TCTGTTTATATCCCTTGAGAGAAAAATATCAAGCAACCTGCCTCAAACGACCCTGTTTTG
GCCGGTGAGCA**AGGACAGCCTCCGATGGATAACTACACAGTG**<u>GCCCCGGACGATGAATAT</u>
<u>GATGTCCTAATCTTAGACGACTACCTGGACAACAGTGGGCCGGACCAAGTTCCGGCCCCC</u>
<u>GAGTTCCTCTCCCCCCAGCAGGTGCTGCAGTTCTGCTGCGCGGTGTTTGCGGTGGGTCTC</u>
<u>TTGGACAACGTGCTGGCGGTGTTTATCTTGGTGAAATACAAAGGACTCAAGAATCTGGGG</u>
<u>AACATCTACTTCCTAAACCTGGCACTTTCAAACCTGTGTTTCCTGCTTCCCCTGCCGTTC</u>
<u>TGGGCCCATACTGCAGCACACGGGGAAAGCCCTGGCAACGGGACCTGTAAAGTTCTTGTC</u>
<u>GGACTCCACTCCTCGGGCTTATACAGCGAGGTGTTTTCCAACATCCTCCTCCTTGTGCAA</u>
<u>GGATACAGGGTGTTTTCCCAAGGGCGACTGGCCTCCATCTTCACGACAGTGTCTTGTGGT</u>
<u>ATTGTTGCGTGCGTCCTGGCATGGG</u>**CCATGGCTACTGCGCTCTCTTTGCCCGAGTCTGTG**
TTTTATGAGCCTCGGATGGAAAGACAGAAACACAAGTGTGCCTTTGGCAAACCTCACTTC
TTGCCAATCGAAGCGCCGCTCTGGAAGTACGTTCTGACGTCAAAAATGATCATCTTGGTA
CTTGCTTTTCCTCTGCTGGTTTTTATAATCTGCTGCAGGCAACTGAGGAGAAGGCAGAGC
TTCAGGGAGAGACAGTACGACCTCCACAAGCCGGCTCTTGTCATAACGGGCGTGTTCCTT
TTGATGTGGGCGCCTTACAACACTGTGCTTTTCCTGTCTGCTTTCCAGGAACACTTGTCC
CTGCAGGATGAGAAGAGCAGCTACCACCTGGACGCAAGTGTTCAGGTCACACAGCTGGTA
GCGACCACCCACTGCTGCGTCAACCCGCTGCTCTATTTGCTTCTTGACCGGAAGGCCTTT
ATGAGATACCTTCGCAGCCTGTTCCCACGGTGCAATGATATCCCCTATCAAAGTAGTGGA
GGCTCTCAGCAAGCGCCGCCAAGGAAAGGTCATGGCAGGCCCATTGAACTGTACAGCAAT
TTGCATCAAAGGCAGGATATAATATAATATAAACACTCAACCTTGTCTCTTCACATTGTT
TTATGCCATTTTTATATGTTTGTATAAAAATAGAATATAAGAGGAAAGGGGGCTGGTTAA
TGAACATTTACTACTAAGCGTTGAATTTGTCTCTGGCACACGGTTAAGTTATTTATAAAG
GTGAATTCCACTCTTCCTCTTTAGCGTTTGTCCGCAGTGTAGGCGTGGGGCGGTGTGCAT
TCCATAAACTGAAGCTTGGTAGTTGTCCAAAGTTACATAATTATAAAGTGGCAAAACTGA
TGTGTAAACTCAGAGATCACTGGCCCCAGAGTCTGTGAGTTATCCAAGCGTGCCAGCCTT
TCAGTAACTGTCAGCTAGCAGTGGTGGCATGTCTGTCATCCCAGCACTCTAGGAGGCGGA
GGCAGGAGGCATCTCACAAATGCAAGTATAACCTGACTTCCCCCAACCTCAATTCTTTTT
TTTTTTGTTGTTGTTGTTCATCAAAATACATGCAATAAATATATATATTTTTAAAGCAAA
AAAAA

FIGURE 16A

Gene Sequence Structure * 463 bp Sequence Deleted 925 bp

Size of full-length cDNA: 2045 bp

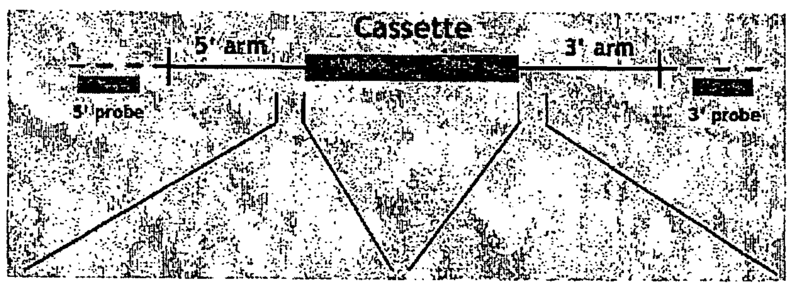


Targeting Vector* (genomic sequence)

Construct Number: 285

Arm Length:
5': 1.3 kb
3': 3.5 kb

Neo



* Not drawn to scale

5'>GACTCTTTCTGTACCAGGCAG TTTCTGTCACATTAGCAGCNCGGT CCCTGGGTGCTTAATGAGGAGATT TGAGTACGGCCGCCCCCGCTGGGG GTGGGGGGTGGGGGCATAGTTACT TTCAGTCTTTGGCTGTCGGATGGA GGGGAATCATGACCTCTGGTTTCC CCACAGGACAGCCTCCGATGGATA ACTACACAGTG<3' (SEQ ID NO:31)	5'>CCATGGCTACTGCGCTCTCTT TGCCCGAGTCTGTGTTTTATGAGC CTCGGATGGAAAGACAGAAACACA AGTGTGCCTTTGGCAAACCTCACT TCTTGCCAATCGAAGCGCCGCTCT GGAAGTACGTTCTGACGTCAAAAA TGATCATCTTGGTACTTGCTTTTC CTCTGCTGGTTTTTATAATCTGCT GCAGGCAACTG<3' (SEQ ID NO:32)

FIGURE 16B

GATTCTTTAGCATGTTGGCTGCTGGAGAAAAGTTCAGGGGCAGGCGACAAGGTCCGGGTGCTGAGCAAGGTGGCT
GGACAAGCCTAGCGGGAGAGAAACCTGAGGCCACGAGCGACCTCCGGAGGAGCCCCTAGACGGGAGTCTTAGCTG
AGACACGGCAGCATGCAACCAGCCCAGATGCGCCTTGCAGAACGAGAGCCTATTTTCAACACTCTTCCCGCAGTC
GGGCCTGCACGCCTTCAGGAACCCACCACCACTCTCCCAATTGTTCCCTGAAATGAACGTTCCTGGCGCGCCCTC
TCGTGGACTGTTGAAGTAACCATCTGTCTCTAGGAAGCCATGGCAACCACGCATGCGCAGGGCCACCAGCCAGTC
TTGGGAAATGACACTCTCCGGGAACATTACGATTACGTGGGGAAACTGGCAGGCAGGCTGCGGGACCCCCCTGAG
GGTGGCACCCTCATCACCACCATCCTCTTCTTGGTCACCTGTAGCTTCATCGTCTTGGAGAACCTGATGGTTTTG
ATTGCCATCTGGAAAAACAATAAATTTCACAACCGCATGTACTTTTTCATCGGCAACTTGGCTCTCTGCGACCTG
CTGGCCGGCATAGCATACAAGGTCAATATTCTTATGTCCGGCAGGAAGACGTTCAGTCTGTCTCCAACAGTGTGG
TTCCTCAGGGAGGGCAGTATGTTCGTAGCCCTGGGCGCATCCACCTGCAGCTTACTGGCCATTGCCATTGAGCGA
CACCTGACCATGATCAAGATGAGGCCATATGATGCCAACAAGAAGCACCGCGTGTTCCTTCTGATTGGGATGTGC
TGGCTAATTGCCTTCTCGCTGGGTGCCCTGCCAATCCTGGGCTGGAACTGCCTGGAGAACTTTCCCGACTGCTCT
ACCATCTTGCCCCTCTACTCCAAGAAATACATCGCCTTCCTCATCAGTATCTTCACCGCCATTCTGGTGACCATC
GTCATCTTGTATGCGCGCATCTACTGCCTGGTCAAGTCCAGCAGCCGCAGGGTGGCCAACCACAACTCTGAGAGA
TCCATGGCCCTTCTGCGGACCGTAGTGATTGTGGTGAGTGTGTTCATTGCCTGTTGGTCCCCACTTTTTATCCTC
TTCCTCATCGACGTGGCCTGCAGGGCAAAGGAGTGCTCCATCCTCTTCAAGAGTCAGTGGTTCATCATGCTGGCT
GTCCTCAACTCGGCCATGAACCCTGTCATCTACACGCTGGCCAGCAAAGAGATGAGGCGCGCCTTCTTCCGGTTG
GTGTGCGGCTGTCTAGTCAAGGGCAAGGGGACCCAGGCCTCACCCATGCAGCCTGCCCTCGACCCAAGCAGAAGT
AAGTCAAGCTCCAGTAACAACAGCAGCCACTCTCCGAAGGTCAAGGAAGACCTGCCCCGCGTGGCTACTTCTTCC
TGCATCATTGACAAAAACAGGTCGTTTCAGAACGGGGTCCTCTGCAAGTGACGGTCTCAACAGGGCAAGCTCTGC
AGCCACACCTATTTATTGCATGCATCACTTCCACGTGGGGCCTTAAGAGATCTTAACTCTGGAGGTCTGCTTGCT
GTGGCCAGACTGTGTGGTATCTCCTGAGGAGGAGACCCAAGGAGTCACCAACGCATTTCAGCCTGGACTTGGATG
TGACTTATCTGTTGCGGGCTCCATCCTTCTGAATGTACCAAGCTGCTGACACGGTGCACTGCCTTCGGATGCCGC
TACCCCTGGTCCAAGGAAAGCAAGGTGGTGAAGAGCTATGCACATGCACATTTGGGATGATTGTCTCCCTTGCGT
TCACACTACATTTATTTTCATAAAA (SEQ ID NO:33)

MATTHAQGHQPVLGNDTLREHYDYVGKLAGRLRDPPEGGTLITTILFLVTCSFIVLENLMVLIAIWKNNKFHNRM
YFFIGNLALCDLLAGIAYKVNILMSGRKTFSLSPTVWFLREGSMFVALGASTCSLLAIAIERHLTMIKMRPYDAN
KKHRVFLLIGMCWLIAFSLGALPILGWNCLENFPDCSTILPLYSKKYIAFLISIFTAILVTIVILYARIYCLVKS
SSRRVANHNSERSMALLRTVVIVVSVFIACWSPLFILFLIDVACRAKECSILFKSQWFIMLAVLNSAMNPVIYTL
ASKEMRRAFFRLVCGCLVKGKGTQASPMQPALDPSRSKSSSSNNSSHSPKVKEDLPRVATSSCIIDKNRSFQNGV
LCK (SEQ ID NO:34)

FIGURE 17

Underlined = deleted in targeting construct

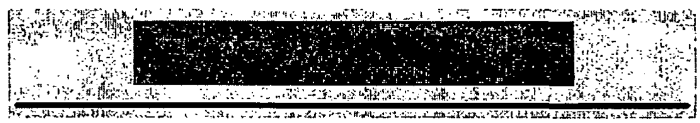
Bold = sequence flanking Neo insert in targeting construct

GATTCTTTAGCATGTTGGCTGCTGGAGAAAAGTTCAGGGGCAGGCGACAAGGTCCGGGTG
CTGAGCAAGGTGGCTGGACAAGCCTAGCGGGAGAGAAACCTGAGGCCACGAGCGACCTCC
GGAGGAGCCCCTAGACGGGAGTCTTAGCTGAGACACGGCAGCATGCAACCAGCCCAGATG
CGCCTTGCAGAACGAGAGCCTATTTTCAACACTCTTCCCGCAGTCGGGCCTGCACGCCTT
CAGGAACCCACCACCACTCTCCCAATTGTTCCCTGAAATGAACGTTCCTGGCGCGCCCTC
TCGTGGACTGTTGAAGTAACCATCTGTCTCTAGGAAGCCATGGCAACCACGCATGCGCAG
GGCCACCAGCCAGTCTTGGGAAATGACACTCTCCGGGAACATTACGATTACGTGGGGAAA
CTGGCAGGCAGGCTGCGGGACCCCCCTGAGGGTGGCACCCTCATCACCACCATCCTCTTC
TTGGTCACCTGTAGCTTCATCGTCTTGG**AGAACCTGATGGTTTTGATTGCCATCTGGAAA**
AACAATAAATTTCACAACCGCATGTACTTTTTCATCGGCAACTTGGCTCTCTGCGACCTG
CTGGCCGGCATAGCATACAAGGTCAATATTCTTATGTCCGGCAGGAAGACGTTCAGTCTG
TCTCCAACAGTGTGGTTCCTCAGGGAGGGCAGTATGTTCGTAGCCCTGGGCGCATCCACC
TGCAGCTTACTGGCCATTGCCATTGAGCGACACCTGACCATGATCAAGATGAGGCCATAT
GATGCCAACAAGAAGCACCGCGTGTTCCTTCTGATTGGGATGTGCTGGCTAATTGCCTTC
TCGCTGGGTGCCCTGCCAATCCTGGGCTGGAACTGCCTGGAGAACTTTCCCGACTGCTCT
ACCATCTTGCCCCTCTACTCCAAGAAATACATCGCCTTCCTCATCAGTATCTTCACCGCC
ATTCTGGTGACCATCGTCATCTTGTATGCGCGCATCTACTGCCTGGTCAAGTCCAGCAGC
CGCAGGGTGGCCAACCACAACTCTGAGAGATCCATGGCCCTTCTGCGGACCGTAGTGATT
GTGGTGAGTGTGTTCATTGCCTGTTGGTCCCCACTTTTTATCCTCTTCCTCATCGACGTG
GCCTGCAGGGCAAAGGAGTGCTCCATCCTCTTCA**AGAGTCAGTGGTTCATCATGCTGGCT**
GTCCTCAACTCGGCCATGAACCCTGTCATCTACACGCTGGCCAGCAAAGAGATGAGGCGC
GCCTTCTTCCGGTTGGTGTGCGGCTGTCTAGTCAAGGGCAAGGGGACCCAGGCCTCACCC
ATGCAGCCTGCCCTCGACCCAAGCAGAAGTAAGTCAAGCTCCAGTAACAACAGCAGCCAC
TCTCCGAAGGTCAAGGAAGACCTGCCCCGCGTGGCTACTTCTTCCTGCATCATTGACAAA
AACAGGTCGTTTCAGAACGGGGTCCTCTGCAAGTGACGGTCTCAACAGGGCAAGCTCTGC
AGCCACACCTATTTATTGCATGCATCACTTCCACGTGGGGCCTTAAGAGATCTTAACTCT
GGAGGTCTGCTTGCTGTGGCCAGACTGTGTGGTATCTCCTGAGGAGGAGACCCAAGGAGT
CACCAACGCATTTCAGCCTGGACTTGGATGTGACTTATCTGTTGCGGGCTCCATCCTTCT
GAATGTACCAAGCTGCTGACACGGTGCACTGCCTTCGGATGCCGCTACCCCTGGTCCAAG
GAAAGCAAGGTGGTGAAGAGCTATGCACATGCACATTTGGGATGATTGTCTCCCTTGCGT
TCACACTACATTTATTTTCATAAAA

FIGURE 18A

Gene Sequence Structure * 1016 bp Sequence Deleted 1174 bp

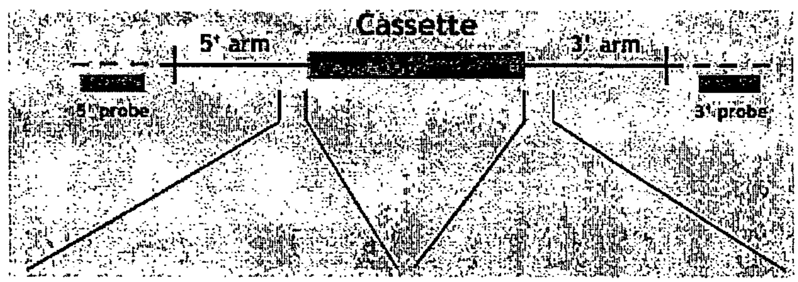
Size of full-length cDNA: 1825 bp



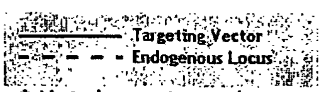
Targeting Vector* (genomic sequence)

Construct Number: 1763

Arm Length:
5': 3.3 kb
3': 2.5 kb



5'>TGGGATGTGCTGGCTAATTGC CTTCTCGCTGGGTGCCCTGCCAAT CCTGGGCTGGAACTGCCTGGAGAA CTTTCCCGACTGCTCTACCATCTT GCCCCTCTACTCCAAGAAATACAT CGCCTTCCTCATCAGTATCTTCAC CGCCATTCTGGTGACCATCGTCAT CTTGTATGCGCGCATCTACTGCCT GGTCAAGTCCA<3' (SEQ ID NO:35)	5'>AGAGTCAGTGGTTCATCATGC TGGCTGTCCTCAACTCGGCCATGA ACCCTGTCATCTACACGCTGGCCA GCAAAGAGATGAGGCGCGCCTTCT TCCGGTTGGTGTGCGGCTGTCTAG TCAAGGGCAAGGGGACCCAGGCCT CACCCATGCAGCCTGCCCTCGACC CAAGCAGAAGTAAGTCAAGCTCCA GTAACAACAGC<3' (SEQ ID NO:36)



* Not drawn to scale

FIGURE 18B

CTGGAACAGCGGCGGCAGCCCTTGGAGGGGACCAGCGCAGCGGGGAGGTGAGAGGTGCTTCCTAGAGGGGAGCGG
ACACCAAGAGAGGTCGCCAAGTGAGGCCTGGGGACGGAGCGCCGGTCGAGTCAGCGCCAAACAGGGAGAGCCCA
TTACGTGACTGTCCCGAGGGCTCGGGCTCCCAGCCCTTGTACTGCGCAAAACTGGGCGGTGAGAGCCGGCTGCCG
GTGCGCTGATGAACTTTGTTTCCTCTAGAGCGTGATCTCGGGCAATGGCAGCAGACCTGGAACCCTGGAATAGCA
CCATCAATGGCACCTGGGAGGGGGACGAACTGGGATACAAGTGTCGCTTCAACGAGGACTTCAAGTACGTGCTGT
TGCCCGTGTCCTATGGCGTGGTGTGCGTGCTCGGGTTGTGCCTGAACGTCGTGGCTCTCTATATCTTCCTATGCC
GCCTCAAAACCTGGAACGCCTCCACCACCTACATGTTTCACCTGGCAGTTTCGGACTCTCTCTACGCAGCGTCCC
TGCCGCTGTTGGTTTATTACTACGCCCGGGGTGACCACTGGCCATTTAGCACGGTGCTCTGCAAGCTGGTGCGTT
TCCTCTTCTACACCAACCTCTACTGCAGCATCCTCTTCCTCACCTGCATCAGCGTGCACCGGTGCCTGGGAGTCC
TGCGCCCTCTGCACTCCCTGCGTTGGGGCCGGGCCCGTTATGCCCGCCGGGTGGCTGCGGTTGTGTGGGTGCTGG
TGCTGGCCTGCCAGGCACCCGTGCTCTACTTCGTCACCACCAGCGTGCGGGGAACCCGGATCACTTGCCATGACA
CCTCGGCCCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCCGTCATGCTGGGTCTGCTTTTTGCTGTGCCCT
TTTCCGTAATCCTGGTCTGTTACGTGCTTATGGCCAGGCGGCTGCTCAAACCGGCTTATGGGACCACAGGAGGTC
TGCCTCGGGCCAAGCGCAAGTCTGTGCGCACCATTGCCTTGGTACTGGCCGTCTTCGCCCTCTGCTTTCTGCCTT
TCCACGTCACGCGCACCCTCTACTACTCCTTCCGATCACTTGACCTCAGCTGCCACACCCTCAACGCCATCAACA
TGGCATATAAGATCACCCGGCCGCTGGCCAGCGCCAACAGTTGTCTTGACCCGGTACTCTACTTCCTGGCAGGGC
AGAGACTTGTCCGCTTTGCCCGAGATGCCAAGCCACCCACGGAGCCTACCCCCAGCCCACAGGCTCGTCGCAAGC
TGGGCCTGCACAGGCCTAACAGAACTGTGAGGAAAGATTTGTCAGTCAGCAGTGACGACTCAAGACGGACAGAGT
CCACACCAGCTGGAAGTGAGACTAAGGACATTCGGCTATAGCACCTCAGTCAGGTGTCAGTTTATGTAGGGGAGC
TGTAGGGGACCAAGGCTTGTTAAAGGGGCTGATATCCTCTGATGCGGATCAACAGCTCTCACTTGGCAGGGGCTC
AGGATACTCATCTTGCTGTCCCAAGAGTAGCAGGACTCCAAAGTCCCAATCATTTGTACGTGTGAGTT[illegible]GAAAC
CATATCTTAAGAAAGGCAGTGCACCCGACTCCTATGCAGACAGCTGCCATTCCTGCCAAGGTACCTAGGCTGAGA
TTCAGCCTCCCAGATCCAAGGGGGAAGCAGGCTCAGAGAGGACACAGCAGAGCCCAGGCTACCTGGATTAGGAGT
AGAATGGGAAAGGTGACTCCACCCCTCTTGTTCCCATGGTGATGTAGGATGGACTGGACCTGGGCACGATGGACT
TAGCTCAGAGGAGGACCCTCAGGCCAAGAGACAAACCTCTGAGAGCTAATGTCAGGTCCCTGGAGACTTCCCCTG
GCTGGCAGCAGGTGGGAACCATAACTTATATGAAAGNCTGTGCACCCATTGGGCTGTGCTCTGGGGTATGACCTG
GGAAGCTGTCACCACCCACACAAAGGNNTGTTATCTCTGCCTCCTNCCTTCCTTCAGTAGTAGGAAGTTTTGGAT
GGGTGTGTGTGTGTGTGTGTGTGTGGGGGGGGTGATCCCTTCTCAATCCCCCAGGTCAACCCTCCTTTCAGCCACTT
TCTCTTTATTCTTCCAAGCATGAGAGGAGCTAAATATG (SEQ ID NO:37)

MAADLEPWNSTINGTWEGDELGYKCRFNEDFKYVLLPVSYGVVCVLGLCLNVVALYIFLCRLKTWNASTTYMFHL
AVSDSLYAASLPLLVYYYARGDHWPFSTVLCKLVRFLFYTNLYCSILFLTCISVHRCLGVLRPLHSLRWGRARYA
RRVAAVVWVLVLACQAPVLYFVTTSVRGTRITCHDTSARELFSHFVAYSSVMLGLLFAVPFSVILVCYVLMARRL
LKPAYGTTGGLPRAKRKSVRTIALVLAVFALCFLPFHVTRTLYYSFRSLDLSCHTLNAINMAYKITRPLASANSC
LDPVLYFLAGQRLVRFARDAKPPTEPTPSPQARRKLGLHRPNRTVRKDLSVSSDDSRRTESTPAGSETKDIRL
(SEQ ID NO:38)

FIGURE 19

<u>Underlined</u> = deleted in targeting construct

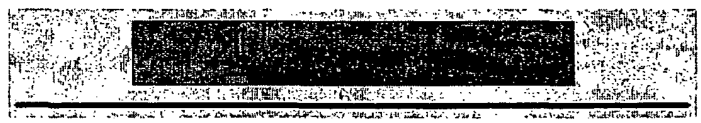
Bold = sequence flanking Neo insert in targeting construct

CTGGAACAGCGGCGGCAGCCCTTGGAGGGGACCAGCGCAGCGGGGAGGTGAGAGGTGCTT
CCTAGAGGGGAGCGGACACCAAGAGAGGTCGCCAAGTGAGGCCTGGGGACGGGAGCGCCG
GTCGAGTCAGCGCCAAACAGGGAGAGCCCATTACGTGACTGTCCCGAGGGCTCGGGCTCC
CAGCCCTTGTACTGCGCAAAACTGGGCGGTGAGAGCCGGCTGCCGGTGCGCTGATGAACT
TTGTTTCCTCTAGAGCGTGATCTCG**GGCAATGGCAGCAGACCTGGAACCCTGGAATAGCA**
CCATCAATGGCACCTGGAGGGGGACGAACTGGGATACAAGTGTCGCTTC<u>AACGAGGACT</u>
<u>TCAAGTACGTGCTGTTGCCCGTGTCCTATGGCGTGGTGTGCGTGCTCGGGTTGTGCCTGA</u>
<u>ACGTCGTGGCTCTCTATATCTTCCTATGCCGCCTCAAAACCTGGAACGCCTCCACCACCT</u>
<u>ACATGTTTCACCTGGCAGTTTCGGACTCTCTCTACGCAGCGTCCCTGCCGCTGTTGGTTT</u>
<u>ATTACTACGCCCGGGGTGACCACTGGCCATTTAGCACGGTGCTCTGCAAGCTGGTGCGTT</u>
<u>TCCTCTTCTACACCAACCTCTACTGCAGCATCCTCT</u>**TCCTCACCTGCATCAGCGTGCACC**
GGTGCCTGGGAGTCCTGCGCCCTCTGCACTCCCTGCGTTGGGGCCGGGCCCGTTATGCCC
GCCGGGTGGCTGCGGTTGTGTGGGTGCTGGTGCTGGCCTGCCAGGCACCCGTGCTCTACT
TCGTCACCACCAGCGTGCGGGGAACCCGGATCACTTGCCATGACACCTCGGCCCGAGAGC
TCTTTAGCCATTTTGTGGCTTACAGCTCCGTCATGCTGGGTCTGCTTTTTGCTGTGCCCT
TTTCCGTAATCCTGGTCTGTTACGTGCTTATGGCCAGGCGGCTGCTCAAACCGGCTTATG
GGACCACAGGAGGTCTGCCTCGGGCCAAGCGCAAGTCTGTGCGCACCATTGCCTTGGTAC
TGGCCGTCTTCGCCCTCTGCTTTCTGCCTTTCCACGTCACGCGCACCCTCTACTACTCCT
TCCGATCACTTGACCTCAGCTGCCACACCCTCAACGCCATCAACATGGCATATAAGATCA
CCCGGCCGCTGGCCAGCGCCAACAGTTGTCTTGACCCGGTACTCTACTTCCTGGCAGGGC
AGAGACTTGTCCGCTTTGCCCGAGATGCCAAGCCACCCACGGAGCCTACCCCCAGCCCAC
AGGCTCGTCGCAAGCTGGGCCTGCACAGGCCTAACAGAACTGTGAGGAAAGATTTGTCAG
TCAGCAGTGACGACTCAAGACGGACAGAGTCCACACCAGCTGGAAGTGAGACTAAGGACA
TTCGGCTATAGCACCTCAGTCAGGTGTCAGTTTATGTAGGGGAGCTGTAGGGGACCAAGG
CTTGTTAAAGGGGCTGATATCCTCTGATGCGGATCAACAGCTCTCACTTGGCAGGGGCTC
AGGATACTCATCTTGCTGTCCCAAGAGTAGCAGGACTCCAAAGTCCCAATCATTTGTACG
TGTGAGTTGGGAAACCATATCTTAAGAAAGGCAGTGCACCCGACTCCTATGCAGACAGCT
GCCATTCCTGCCAAGGTACCTAGGCTGAGATTCAGCCTCCCAGATCCAAGGGGGAAGCAG
GCTCAGAGAGGACACAGCAGAGCCCAGGCTACCTGGATTAGGAGTAGAATGGGAAAGGTG
ACTCCACCCCTCTTGTTCCCATGGTGATGTAGGATGGACTGGACCTGGGCACGATGGACT
TAGCTCAGAGGAGGACCCTCAGGCCAAGAGACAAACCTCTGAGAGCTAATGTCAGGTCCC
TGGAGACTTCCCCTGGCTGGCAGCAGGTGGGAACCATAACTTATATGAAAGNCTGTGCAC
CCATTGGGCTGTGCTCTGGGGTATGACCTGGGAAGCTGTCACCACCCACACAAAGGNNTG
TTATCTCTGCCTCCTNCCTTCCTTCAGTAGTAGGAAGTTTTGGATGGGTGTGTGTGTGTG
TGTGTGTGGGGGGGGTGATCCCTTCTCAATCCCCCAGGTCAACCCTCCTTTCAGCCACTT
TCTCTTTATTCTTCCAAGCATGAGAGGAGCTAAATATG

FIGURE 20A

Gene Sequence Structure * 351 bp Sequence Deleted 636 bp

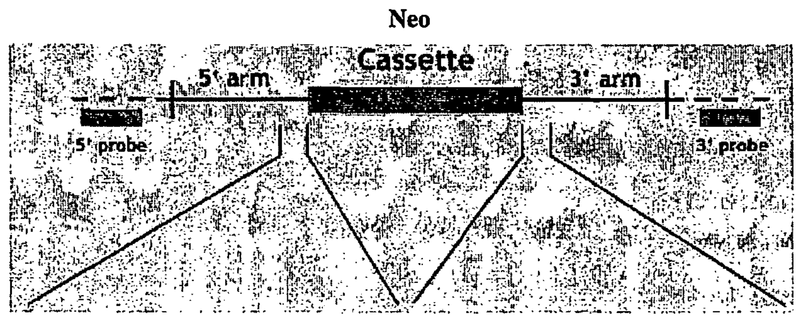
Size of full-length cDNA: 2138 bp

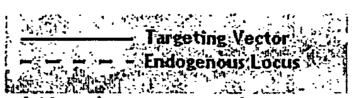


Targeting Vector* (genomic sequence)

Construct Number: 2085

Arm Length:
5': 5 kb
3': 2 kb





* Not drawn to scale

5'>AAGCAAGCGATCTATCACCAA GCTCTACCTCCAGCCTTAGGCAAA TCTTAGCAAGTCCCCTCCCCCTCT GAGCCCAGACATGACAGCTTTCTT ACCACACCTCCCTTTCCTGCAGGG CAATGGCAGCGGACCTGGAACCCT GGAATAGCACCATCAATGGCACCT GGGAGGGGGACGAACTGGGATACA AGTGTCGCTTC<3' (SEQ ID NO:39)	5'>TCCTCACCTGCATCAGCGTGC ACCGGTGCCTGGGAGTCCTGCGCC CTCTGCACTCCCTGCGTTGGGGCC GGGCCCGTTATGCCCGCCGGGTGG CTGCGGTTGTGTGGGTGCTGGTGC TGGCCTGCCAGGCGCCCGTGCTCT ACTTCGTCACCACCAGCGTGCGGG GAACCCGGATCACTTGCCATGACA CCTCGGCCCGA<3' (SEQ ID NO:40)

FIGURE 20B

CTGCGGGTCAAGCTGACAGTCAGATATAGAACGGTTACCACTCAAAGAAGAATATGGCTATTCTTGGGCCTTTGC
TGGCTAGTTTCCTTTCTGGTGGGGCTGACCCCCATGTTTGGCTGGAATAGAAAAGCAACCTTAGCGAGCTCTCAA
AATAGCAGCACTCTTTTATGCCACTTCCGTTCCGTGGTCAGTTTGGATTACATGGTCTTCTTCAGCTTCGTCACC
TGGATCCTCGTCCCCCTGGTTGTCATGTGTGTCATCTACCTAGACATCTTCTACATCATCCGAAATAAGCTCAGT
CAAAACCTGTCTGGCTTCAGAGAGACGCGTGCATTTTATGGACGGGAGTTCAAGACAGCTAAGTCCCTGTTTCTG
GTTCTCTTCTTG (SEQ ID NO:41)

LRVKLTVRYRTVTTQRRIWLFLGLCWLVSFLVGLTPMFGWNRKATLASSQNSSTLLCHFRSVVSLDYMVFFSFVT
WILVPLVVMCVIYLDIFYIIRNKLSQNLSGFRETRAFYGREFKTAKSLFLVLFL (SEQ ID NO:42)

FIGURE 21

Underlined = deleted in targeting construct

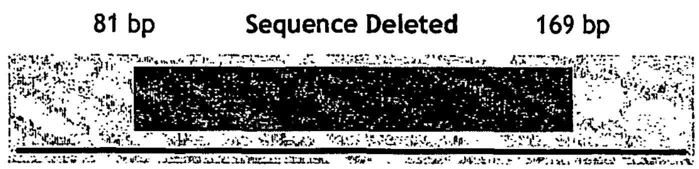
Bold = sequence flanking Neo insert in targeting construct

CTGCGGGTCAAGCTGACAGTCAG**ATATAGAACGGTTACCACTCAAAGAAGAATATGGCTA**
TTCTTGGGCCTTTGCTGGCTAGTTTCCTTTCTGGTGGGGCTGACCCCCATGTTTGGCTGG
AATAGAAAAGCAACCTTAGCGAGCTCTCAAAATAGCAGCACTCTTTTAT**GCCACTTCCGT**
TCCGTGGTCAGTTTGGATTACATGGTCTTCTTCAGCTTCGTCACCTGGATCCTCGTCCCC
CTGGTTGTCATGTGTGTCATCTACCTAGACATCTTCTACATCATCCGAAATAAGCTCAGT
CAAAACCTGTCTGGCTTCAGAGAGACGCGTGCATTTTATGGACGGGAGTTCAAGACAGCT
AAGTCCCTGTTTCTGGTTCTCTTCTTG

FIGURE 22A

Gene Sequence Structure *

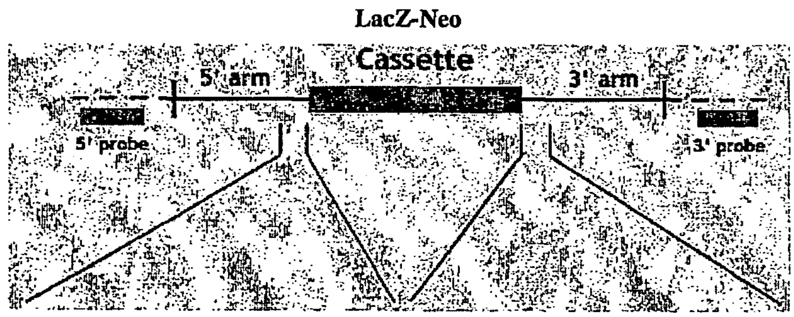
Size of partial cDNA: 387 bp

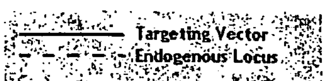


Targeting Vector*
(genomic sequence)

Construct Number: 2827

Arm Length:
5': 1.6 kb
3': 5 kb





* Not drawn to scale

5'>GGATTTAGTCAGAGTGACATT AGGCCATGGCTATTTCACCTCAGC AGATTAGGGGATTGGGAGAGCTGA CTTTGTCTTTCTTGTTTTCTCCCA GGGAGTCATTTCATTGTTGTGAAC TCACCATTTTTCCCCCATTCAAAC AGATATAGAACGGTTACCACTCAA AGAAGAATATGGCTATTCTTGGGC CTTTGCTGGCT<3' (SEQ ID NO:43)	5'>GCCACTTCCGTTCCGTGGTCA GTTTGGATTACATGGTCTTCTTCA GCTTCGTCACCTGGATCCTCGTCC CCCTGGTTGTCATGTGTGTCATCT ACCTAGACATCTTCTACATCATCC GAAATAAGCTCAGTCAAAACCTGT CTGGCTTCAGAGAGACGCGTGCAT TTTATGGACGGGAGTTCAAGACAG CTAAGTCCCTG<3' (SEQ ID NO:44)

FIGURE 22B

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/003793 A3

(51) International Patent Classification[7]: A01K 67/027, C07K 14/705, 14/72, 14/715, A61K 49/00, G01N 33/50

(21) International Application Number: PCT/US01/21923

(22) International Filing Date: 10 July 2001 (10.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/217,058	10 July 2000 (10.07.2000)	US
60/217,179	10 July 2000 (10.07.2000)	US
60/217,223	10 July 2000 (10.07.2000)	US
60/217,253	10 July 2000 (10.07.2000)	US
60/217,255	10 July 2000 (10.07.2000)	US
60/217,256	10 July 2000 (10.07.2000)	US
60/217,257	10 July 2000 (10.07.2000)	US
60/217,347	11 July 2000 (11.07.2000)	US
60/217,629	11 July 2000 (11.07.2000)	US
60/217,537	12 July 2000 (12.07.2000)	US
60/218,069	12 July 2000 (12.07.2000)	US
60/218,074	12 July 2000 (12.07.2000)	US
60/218,358	12 July 2000 (12.07.2000)	US
60/221,483	27 July 2000 (27.07.2000)	US
60/223,120	7 August 2000 (07.08.2000)	US
60/223,122	7 August 2000 (07.08.2000)	US
60/243,958	26 October 2000 (26.10.2000)	US
60/249,408	15 November 2000 (15.11.2000)	US
60/252,299	20 November 2000 (20.11.2000)	US
60/262,113	16 January 2001 (16.01.2001)	US
60/262,205	16 January 2001 (16.01.2001)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/217,179 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,058 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,223 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,253 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,255 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,256 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,257 (CON)
Filed on	10 July 2000 (10.07.2000)
US	60/217,347 (CON)
Filed on	11 July 2000 (11.07.2000)
US	60/217,629 (CON)
Filed on	11 July 2000 (11.07.2000)
US	60/217,537 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,069 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,074 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/218,358 (CON)
Filed on	12 July 2000 (12.07.2000)
US	60/221,483 (CON)
Filed on	27 July 2000 (27.07.2000)
US	60/223,120 (CON)
Filed on	7 August 2000 (07.08.2000)
US	60/223,122 (CON)
Filed on	7 August 2000 (07.08.2000)
US	60/243,958 (CON)
Filed on	26 October 2000 (26.10.2000)
US	60/249,408 (CON)
Filed on	15 November 2000 (15.11.2000)
US	60/252,299 (CON)
Filed on	20 November 2000 (20.11.2000)
US	60/262,113 (CON)
Filed on	16 January 2001 (16.01.2001)
US	60/262,205 (CON)
Filed on	16 January 2001 (16.01.2001)

(71) **Applicant** *(for all designated States except US)*: **DELTAGEN, INC.** [US/US]; 740 Bay Road, Redwood City, CA 94063 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **ALLEN, Keith, D.** [US/US]; 224 Custer Trail, Cary, NC 27513 (US). **BRENNAN, Thomas, J.** [US/US]; 325 Rockwood Drive, South San Francisco, CA 94080 (US).

(74) **Agents: BURKE, John, E.** et al.; Deltagen, Inc., 740 Bay Road, Redwood City, CA 94063 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,

[Continued on next page]

(54) **Title:** TRANSGENIC MICE CONTAINING TARGETED GENE DISRUPTIONS

(57) **Abstract:** The present invention relates to transgenic animals, as well as compositions and methods relating to the characterization of gene function. Specifically, the present invention provides transgenic mice comprising mutations in a GPCR gene. Such transgenic mice are useful as models for disease and for identifying agents that modulate gene expression and gene function, and as potential treatments for various disease states and disease conditions.

WO 02/003793 A3

LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) **Date of publication of the international search report:** 13 February 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/21923

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A01K67/027 C07K14/705 C07K14/72 C07K14/715 A61K49/00
G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 A01K C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, SEQUENCE SEARCH, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 10068 A (UNIV OREGON HEALTH SCIENCES ;FAN WEI (US); LU DONGSI (US); BOSTON) 12 March 1998 (1998-03-12) claim 1 ---	17-50
A	WO 00 27862 A (HUANG RUEY RUEY C ;PLOEG LEONARDUS H T V D (US); FONG TUNG M (US);) 18 May 2000 (2000-05-18) claim 35 ---	17-50
P,X	WO 01 33954 A (PLOEG LEONARDUS H T V D ;CHEN AIRU S (US); CHEN HOWARD Y (US); MER) 17 May 2001 (2001-05-17) figure 6 ---	1-50

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 July 2002	25.10.2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lonnoy, O

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	BUTLER A A ET AL: "A unique metabolic syndrome causes obesity in the melanocortin-3 receptor-deficient mouse." ENDOCRINOLOGY. UNITED STATES SEP 2000, vol. 141, no. 9, September 2000 (2000-09), pages 3518-3521, XP002205109 ISSN: 0013-7227 the whole document ---	1-50
A	WO 97 47316 A (MILLENNIUM PHARM INC) 18 December 1997 (1997-12-18) figures 2,3 ---	1-16
A	WO 98 56914 A (UNIV OREGON HEALTH SCIENCES) 17 December 1998 (1998-12-17) -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/21923

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.: 16, 32, 51 and 52
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
17-52 totally and 1-16 partially

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 16, 32, 51 and 52

Present claims 16, 32, 51 and 52 each relate to a compound defined by reference to a desirable characteristic or property, namely that it ameliorates a phenotype associated with a disruption in a nuclear hormone receptor gene, that it modulates nuclear hormone receptor expression, that it modulates a behavior associated with a disruption in a nuclear hormone recptor gene, or that it modulates nuclear hormone receptor gene function. The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for no such compound. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define a compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search can be carried out for such speculative claims, the wording of which is a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 17-52 (all totally) and 1-16 (all partially)

   A targeting construct comprising a first polynucleotide sequence homologous to a target gene, a second polynucleotide sequence homologous to the target gene, and a selectable marker; a method of producing said targeting construct; a cell comprising a disrupted said target gene; a non-human transgenic animal comprising a disruption in said target gene; a cell derived from said transgenic animal; a method of producing a transgenic mouse comprising a disruption in said target gene; a method of identifying an agent that modulates the expression of said target gene comprising the use of said transgenic animal or of said cell; a method of identifying an agent that modulates the function of said target gene comprising the use of said transgenic animal or of said cell; Said subject-matter wherein said target gene is a melanocortin-3 receptor gene.

2. Claims: 53-78 (all totally) and 1-16 (all partially)

   As for subject 1, but wherein said target gene is a 5HT-2B gene.

3. Claims: 79-108 (all totally) and 1-16 (all partially)

   As for subject 1, but wherein said target gene is a chemokine receptor 9A gene.

4. Claims: 109-143 (all totally) and 1-16 (all partially)

   As for subject 1, but wherein said target gene is a glucocorticoid-induced receptor gene.

5. Claims: 1-16 (all partially)

   As for subject 1, but wherein said target gene is an orphan GPR10 gene.

6. Claims: 1-16 (all partially)

   As for subject 1, but wherein said target gene is an orphan GPR14 gene.

7. Claims: 1-16 (all partially)

   As for subject 1, but wherein said target gene is an orphan

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

GPR15 gene.

8. Claims: 1-16 (all partially)

As for subject 1, but wherein said target gene is a beta chemokine receptor (E01) gene.

9. Claims: 1-16 (all partially)

As for subject 1, but wherein said target gene is an endothelial differentiation GPCR3 (EDG3) gene.

10. Claims: 1-16 (all partially)

As for subject 1, but wherein said target gene is an ATP receptor P2U1 gene.

11. Claims: 1-16 (all partially)

As for subject 1, but wherein said target gene is a adenosine 3 receptor gene.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/21923

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 9810068	A	12-03-1998	US	6100048	A	08-08-2000
			AU	719954	B2	18-05-2000
			AU	4181297	A	26-03-1998
			EP	0935655	A1	18-08-1999
			JP	2001510984	T	07-08-2001
			WO	9810068	A2	12-03-1998
			US	6278038	B1	21-08-2001
WO 0027862	A	18-05-2000	EP	1129104	A1	05-09-2001
			WO	0027862	A1	18-05-2000
WO 0133954	A	17-05-2001	AU	1758401	A	06-06-2001
			EP	1241933	A1	25-09-2002
			WO	0133954	A1	17-05-2001
			EP	1241934	A1	25-09-2002
			WO	0133956	A1	17-05-2001
WO 9747316	A	18-12-1997	US	5908609	A	01-06-1999
			US	5932779	A	03-08-1999
			AU	723135	B2	17-08-2000
			AU	3383697	A	07-01-1998
			BR	9709684	A	09-05-2000
			CA	2257857	A1	18-12-1997
			CN	1227496	A	01-09-1999
			EP	0915706	A1	19-05-1999
			JP	2002514041	T	14-05-2002
			TR	9802572	T2	22-03-1999
			WO	9747316	A1	18-12-1997
			US	6287763	B1	11-09-2001
WO 9856914	A	17-12-1998	AU	7959598	A	30-12-1998
			US	6278038	B1	21-08-2001
			WO	9856914	A1	17-12-1998